

SEQUENCE LISTING

```

#          SEQUENCE LISTING          - - - - (1) GENERAL INFORMATION:
- -      (iii) NUMBER OF SEQUENCES: 26
- -      - - (2) INFORMATION FOR SEQ ID NO:1:
- -          (i) SEQUENCE CHARACTERISTICS:
- -              (A) LENGTH: 1057 base - #pairs
- -              (B) TYPE: nucleic acid              (C) STRANDEDNESS: double
- -              (D) TOPOLOGY: linear          - -      (ix) FEATURE:
- -              (A) NAME/KEY: CDS              (B) LOCATION: 124..893
- -              (ix) FEATURE:                  (A) NAME/KEY: misc.sub.-- - #feature
- -              (B) LOCATION: 1..1057
- -              (D) OTHER INFORMATION: - #/note= "product = Arabidopsis
- -                  thaliana - #AP1."
- -              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
- -      CTTTCCAATT GGTTCATACC AAAGTCTGAG CTCTTCTTTA TATCTCTCTT GT -
#AGTTTCTT      60
- -      ATTGGGGGTC TTTGTTTTGT TTGGTTCTTT TAGAGTAAGA AGTTTCTTAA AA -
#AAGGATCA      120
- -      AAA ATG GGA AGG GGT AGG GTT CAA TTG AAG AG - #G ATA GAG AAC AAG
ATC      168
      Met Gly Arg Gly Arg Val Gln Leu - #Lys Arg Ile Glu Asn Lys Ile
      1          - # 5          - # 10          - # 15
- -      AAT AGA CAA GTG ACA TTC TCG AAA AGA AGA GC - #T GGT CTT TTG AAG AAA
      216
      Asn Arg Gln Val Thr Phe Ser Lys Arg Arg Al - #a Gly Leu Leu Lys Lys
      20 - #          25 - #          30
- -      GCT CAT GAG ATC TCT GTT CTC TGT GAT GCT GA - #A GTT GCT CTT GTT GTC
      264
      Ala His Glu Ile Ser Val Leu Cys Asp Ala Gl - #u Val Ala Leu Val Val
      35          - #          40          - #          45
- -      TTC TCC CAT AAG GGG AAA CTC TTC GAA TAC TC - #C ACT GAT TCT TGT ATG
      312
      Phe Ser His Lys Gly Lys Leu Phe Glu Tyr Se - #r Thr Asp Ser Cys Met
      50          - #          55          - #          60
- -      GAG AAG ATA CTT GAA CGC TAT GAG AGG TAC TC - #T TAC GCC GAA AGA CAG
      360
      Glu Lys Ile Leu Glu Arg Tyr Glu Arg Tyr Se - #r Tyr Ala Glu Arg Gln
      65          - #          70          - #          75
- -      CTT ATT GCA CCT GAG TCC GAC GTC AAT ACA AA - #C TGG TCG ATG GAG TAT
      408
      Leu Ile Ala Pro Glu Ser Asp Val Asn Thr As - #n Trp Ser Met Glu Tyr
      80          - # 85          - # 90          - # 95
- -      AAC AGG CTT AAG GCT AAG ATT GAG CTT TTG GA - #G AGA AAC CAG AGG CAT
      456
      Asn Arg Leu Lys Ala Lys Ile Glu Leu Leu Gl - #u Arg Asn Gln Arg His
      100 - #          105 - #          110
- -      TAT CTT GGG GAA GAC TTG CAA GCA ATG AGC CC - #T AAA GAG CTT CAG AAT
      504
      Tyr Leu Gly Glu Asp Leu Gln Ala Met Ser Pr - #o Lys Glu Leu Gln Asn
      115          - #          120          - #          125

```

- - CTG GAG CAG CAG CTT GAC ACT GCT CTT AAG CA - #C ATC CGC ACT AGA AAA
 552
 Leu Glu Gln Gln Leu Asp Thr Ala Leu Lys Hi - #s Ile Arg Thr Arg Lys
 130 - # 135 - # 140
 5 - - AAC CAA CTT ATG TAC GAG TCC ATC AAT GAG CT - #C CAA AAA AAG GAG AAG
 600
 Asn Gln Leu Met Tyr Glu Ser Ile Asn Glu Le - #u Gln Lys Lys Glu Lys
 145 - # 150 - # 155
 10 - - GCC ATA CAG GAG CAA AAC AGC ATG CTT TCT AA - #A CAG ATC AAG GAG AGG
 648
 Ala Ile Gln Glu Gln Asn Ser Met Leu Ser Ly - #s Gln Ile Lys Glu Arg
 160 1 - #65 1 - #70 1 -
 #75
 - - GAA AAA ATT CTT AGG GCT CAA CAG GAG CAG TG - #G GAT CAG CAG AAC
 15 CAA 696
 Glu Lys Ile Leu Arg Ala Gln Gln Glu Gln Tr - #p Asp Gln Gln Asn Gln
 180 - # 185 - # 190
 - - GGC CAC AAT ATG CCT CCC CCT CTG CCA CCG CA - #G CAG CAC CAA ATC CAG
 744
 20 Gly His Asn Met Pro Pro Pro Leu Pro Pro Gl - #n Gln His Gln Ile Gln
 195 - # 200 - # 205
 - - CAT CCT TAC ATG CTC TCT CAT CAG CCA TCT CC - #T TTT CTC AAC ATG GGT
 792
 25 His Pro Tyr Met Leu Ser His Gln Pro Ser Pr - #o Phe Leu Asn Met Gly
 210 - # 215 - # 220
 - - GGT CTG TAT CAA GAA GAT GAT CCA ATG GCA AT - #G AGG AGG AAT GAT CTC
 840
 Gly Leu Tyr Gln Glu Asp Asp Pro Met Ala Me - #t Arg Arg Asn Asp Leu
 225 - # 230 - # 235
 30 - - GAA CTG ACT CTT GAA CCC GTT TAC AAC TGC AA - #C CTT GGC TGC TTC GCC
 888
 Glu Leu Thr Leu Glu Pro Val Tyr Asn Cys As - #n Leu Gly Cys Phe Ala
 240 2 - #45 2 - #50 2 -
 #55
 35 - - GCA TG AAGCATTTCC ATATATATAT TTGTAATCGT CAACAATAAA AAC - #AGTTTGC
 943 Ala
 - - CACATACATA TAAATAGTGG CTAGGCTCTT TTCATCCAAT TAATATATTT TG -
 #GCAAATGT 1003
 - - TCGATGTTCT TATATCATCA TATATAAATT AGCAGGCTCC TTTCTTTTTT TG - #TA
 40 1057 - - - - (2) INFORMATION FOR SEQ ID NO:2:
 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 256 amino - #acids (B) TYPE: amino acid
 (D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: protein
 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
 45 - - Met Gly Arg Gly Arg Val Gln Leu Lys Arg Il - #e Glu Asn Lys Ile Asn
 1 5 - # 10 - # 15
 - - Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gl - #y Leu Leu Lys Lys Ala
 20 - # 25 - # 30
 - - His Glu Ile Ser Val Leu Cys Asp Ala Glu Va - #l Ala Leu Val Val Phe
 50 35 - # 40 - # 45
 - - Ser His Lys Gly Lys Leu Phe Glu Tyr Ser Th - #r Asp Ser Cys Met Glu

197
 Cys Asp Ala Glu Val Ala Leu Val Val Phe Se - #r His Lys Gly Lys Leu
 40 - # 45 - # 50
 - - TTT GAA TAC TCC ACT GAT TCT TGT ATG GAG AA - #G ATA CTT GAA CGC TAT
 245
 Phe Glu Tyr Ser Thr Asp Ser Cys Met Glu Ly - #s Ile Leu Glu Arg Tyr
 55 - # 60 - # 65 - # 70
 - - GAG AGA TAC TCT TAC GCC GAG AGA CAG CTT AT - #A GCA CCT GAG TCC GAC
 293
 Glu Arg Tyr Ser Tyr Ala Glu Arg Gln Leu Il - #e Ala Pro Glu Ser Asp
 75 - # 80 - # 85
 - - TCC AAT ACG AAC TGG TCG ATG GAG TAT AAT AG - #G CTT AAG GCT AAG ATT
 341
 Ser Asn Thr Asn Trp Ser Met Glu Tyr Asn Ar - #g Leu Lys Ala Lys Ile
 90 - # 95 - # 100
 - - GAG CTT TTG GAG AGA AAC CAG AGG CAC TAT CT - #T GGG GAA GAC TTG CAA
 389
 Glu Leu Leu Glu Arg Asn Gln Arg His Tyr Le - #u Gly Glu Asp Leu Gln
 105 - # 110 - # 115
 - - GCA ATG AGC CCT AAG GAA CTC CAG AAT CTA GA - #G CAA CAG CTT GAT ACT
 437
 Ala Met Ser Pro Lys Glu Leu Gln Asn Leu Gl - #u Gln Gln Leu Asp Thr
 120 - # 125 - # 130
 - - GCT CTT AAG CAC ATC CGC TCT AGA AAA AAC CA - #A CTT ATG TAC GAC TCC
 485
 Ala Leu Lys His Ile Arg Ser Arg Lys Asn Gl - #n Leu Met Tyr Asp Ser
 135 1 - #40 1 - #45 1 -
 #50
 - - ATC AAT GAG CTC CAA AGA AAG GAG AAA GCC AT - #A CAG GAA CAA AAC
 AGC 533
 Ile Asn Glu Leu Gln Arg Lys Glu Lys Ala Il - #e Gln Glu Gln Asn Ser
 155 - # 160 - # 165
 - - ATG CTT TCC AAG CAG ATT AAG GAG AGG GAA AA - #C GTT CTT AGG GCG CAA
 581
 Met Leu Ser Lys Gln Ile Lys Glu Arg Glu As - #n Val Leu Arg Ala Gln
 170 - # 175 - # 180
 - - CAA GAG CAA TGG GAC GAG CAG AAC CAT GGC CA - #T AAT ATG CCT CCG CCT
 629
 Gln Glu Gln Trp Asp Glu Gln Asn His Gly Hi - #s Asn Met Pro Pro Pro
 185 - # 190 - # 195
 - - CCA CCC CCG CAG CAG CAT CAA ATC CAG CAT CC - #T TAC ATG CTC TCT CAT
 677
 Pro Pro Pro Gln Gln His Gln Ile Gln His Pr - #o Tyr Met Leu Ser His
 200 - # 205 - # 210
 - - CAG CCA TCT CCT TTT CTC AAC ATG GGG GGG CT - #G TAT CAA GAA GAA GAT
 725
 Gln Pro Ser Pro Phe Leu Asn Met Gly Gly Le - #u Tyr Gln Glu Glu Asp
 215 2 - #20 2 - #25 2 -
 #30
 - - CAA ATG GCA ATG AGG AGG AAC GAT CTC GAT CT - #G TCT CTT GAA CCC
 GGT 773

Gln Met Ala Met Arg Arg Asn Asp Leu Asp Le - #u Ser Leu Glu Pro Gly
235 - # 240 - # 245
- - TAT AAC TGC AAT CTC GGC TGC - # - #
794 Tyr Asn Cys Asn Leu Gly Cys 250
- - - (2) INFORMATION FOR SEQ ID NO:4:
- - (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 253 amino - #acids (B) TYPE: amino acid
(D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: protein
- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
- - Met Gly Arg Gly Arg Val Gln Leu Lys Arg Il - #e Glu Asn Lys Ile Asn
1 5 - # 10 - # 15
- - Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gl - #y Leu Met Lys Lys Ala
20 - # 25 - # 30
- - His Glu Ile Ser Val Leu Cys Asp Ala Glu Va - #l Ala Leu Val Val Phe
35 - # 40 - # 45
- - Ser His Lys Gly Lys Leu Phe Glu Tyr Ser Th - #r Asp Ser Cys Met Glu
50 - # 55 - # 60
- - Lys Ile Leu Glu Arg Tyr Glu Arg Tyr Ser Ty - #r Ala Glu Arg Gln Leu
65 - # 70 - # 75 - # 80
- - Ile Ala Pro Glu Ser Asp Ser Asn Thr Asn Tr - #p Ser Met Glu Tyr Asn
85 - # 90 - # 95
- - Arg Leu Lys Ala Lys Ile Glu Leu Leu Glu Ar - #g Asn Gln Arg His Tyr
100 - # 105 - # 110
- - Leu Gly Glu Asp Leu Gln Ala Met Ser Pro Ly - #s Glu Leu Gln Asn Leu
115 - # 120 - # 125
- - Glu Gln Gln Leu Asp Thr Ala Leu Lys His Il - #e Arg Ser Arg Lys Asn
130 - # 135 - # 140
- - Gln Leu Met Tyr Asp Ser Ile Asn Glu Leu Gl - #n Arg Lys Glu Lys Ala
145 1 - #50 1 - #55 1 -
#60
- - Ile Gln Glu Gln Asn Ser Met Leu Ser Lys Gl - #n Ile Lys Glu Arg
Glu
165 - # 170 - # 175
- - Asn Val Leu Arg Ala Gln Gln Glu Gln Trp As - #p Glu Gln Asn His Gly
180 - # 185 - # 190
- - His Asn Met Pro Pro Pro Pro Pro Gln Gl - #n His Gln Ile Gln His
195 - # 200 - # 205
- - Pro Tyr Met Leu Ser His Gln Pro Ser Pro Ph - #e Leu Asn Met Gly Gly
210 - # 215 - # 220
- - Leu Tyr Gln Glu Glu Asp Gln Met Ala Met Ar - #g Arg Asn Asp Leu Asp
225 2 - #30 2 - #35 2 -
#40 - - Leu Ser Leu Glu Pro Gly Tyr Asn Cys Asn Le - #u Gly Cys
245 - # 250
- - - (2) INFORMATION FOR SEQ ID NO:5:
- - (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 768 base - #pairs (C) STRANDEDNESS: double
(B) TYPE: nucleic acid (D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: cDNA
- - (ix) FEATURE: (A) NAME/KEY: CDS
(B) LOCATION: 1..766 - - (ix) FEATURE:

(A) NAME/KEY: misc.sub.-- - #feature
 (B) LOCATION: 1..768
 (D) OTHER INFORMATION: - #/note= "product = Brassica oleracea
 var. botr - #ytis APl."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
 - - ATG GGA AGG GGT AGG GTT CAG TTG AAG AGG AT - #A GAA AAC AAG ATC AAT
 48
 Met Gly Arg Gly Arg Val Gln Leu Lys Arg Il - #e Glu Asn Lys Ile Asn
 1 5 - # 10 - # 15
 - - AGA CAA GTG ACA TTC TCG AAA AGA AGA GCT GG - #T CTT ATG AAG AAA GCT
 96
 Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gl - #y Leu Met Lys Lys Ala
 20 - # 25 - # 30
 - - CAT GAG ATC TCT GTT CTG TGT GAT GCT GAA GT - #T GCG CTT GTT GTC TTC
 144
 His Glu Ile Ser Val Leu Cys Asp Ala Glu Va - #l Ala Leu Val Val Phe
 35 - # 40 - # 45
 - - TCC CAT AAG GGG AAA CTC TTT GAA TAC CCC AC - #T GAT TCT TGT ATG GAG
 192
 Ser His Lys Gly Lys Leu Phe Glu Tyr Pro Th - #r Asp Ser Cys Met Glu
 50 - # 55 - # 60
 - - GAG ATA CTT GAA CGC TAT GAG AGA TAC TCT TA - #C GCC GAG AGA CAG CTT
 240
 Glu Ile Leu Glu Arg Tyr Glu Arg Tyr Ser Ty - #r Ala Glu Arg Gln Leu
 65 - # 70 - # 75 - # 80
 - - ATA GCA CCT GAG TCC GAC TCC AAT ACG AAC TG - #G TCG ATG GAG TAT AAT
 288
 Ile Ala Pro Glu Ser Asp Ser Asn Thr Asn Tr - #p Ser Met Glu Tyr Asn
 85 - # 90 - # 95
 - - AGG CTT AAG GCT AAG ATT GAG CTT TTG GAG AG - #A AAC CAG AGG CAC TAT
 336
 Arg Leu Lys Ala Lys Ile Glu Leu Leu Glu Ar - #g Asn Gln Arg His Tyr
 100 - # 105 - # 110
 - - CTT GGG GAA GAC TTG CAA GCA ATG AGC CCT AA - #G GAA CTC CAG AAT CTA
 384
 Leu Gly Glu Asp Leu Gln Ala Met Ser Pro Ly - #s Glu Leu Gln Asn Leu
 115 - # 120 - # 125
 - - GAG CAA CAG CTT GAT ACT GCT CTT AAG CAC AT - #C CGC TCT AGA AAA AAC
 432
 Glu Gln Gln Leu Asp Thr Ala Leu Lys His Il - #e Arg Ser Arg Lys Asn
 130 - # 135 - # 140
 - - CAA CTT ATG TAC GAC TCC ATC AAT GAG CTC CA - #A AGA AAG GAG AAA GCC
 480
 Gln Leu Met Tyr Asp Ser Ile Asn Glu Leu Gl - #n Arg Lys Glu Lys Ala
 145 1 - #50 1 - #55 1 -
 #60
 - - ATA CAG GAA CAA AAC AGC ATG CTT TCC AAG CA - #G ATT AAG GAG AGG
 GAA 528
 Ile Gln Glu Gln Asn Ser Met Leu Ser Lys Gl - #n Ile Lys Glu Arg Glu
 165 - # 170 - # 175
 - - AAC GTT CTT AGG GCG CAA CAA GAG CAA TGG GA - #C GAG CAG AAC CAT GGC

576

Asn Val Leu Arg Ala Gln Gln Glu Gln Trp As - #p Glu Gln Asn His Gly
 180 - # 185 - # 190
 - - CAT AAT ATG CCT CCG CCT CCA CCC CCG CAG CA - #G CAT CAA ATC CAG CAT
 624

His Asn Met Pro Pro Pro Pro Pro Gln Gl - #n His Gln Ile Gln His
 195 - # 200 - # 205
 - - CCT TAC ATG CTC TCT CAT CAG CCA TCT CCT TT - #T CTC AAC ATG GGA GGG
 672

Pro Tyr Met Leu Ser His Gln Pro Ser Pro Ph - #e Leu Asn Met Gly Gly
 210 - # 215 - # 220
 - - CTG TAT CAA GAA GAA GAT CAA ATG GCA ATG AG - #G AGG AAC GAT CTC GAT
 720

Leu Tyr Gln Glu Glu Asp Gln Met Ala Met Ar - #g Arg Asn Asp Leu Asp
 225 2 - #30 2 - #35 2 -
 #40

- - CTG TCT CTT GAA CCC GTT TAC AAC TGC AAC CT - #T GGC CGT CGC TGC T
 766

Leu Ser Leu Glu Pro Val Tyr Asn Cys Asn Le - #u Gly Arg Arg Cys
 245 - # 250 - # 255

- - GA - # - # - #

768 - - - - (2) INFORMATION FOR SEQ ID NO:6:

- - (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 255 amino - #acids (B) TYPE: amino acid

(D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: protein

- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

- - Met Gly Arg Gly Arg Val Gln Leu Lys Arg Il - #e Glu Asn Lys Ile Asn
 1 5 - # 10 - # 15

- - Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gl - #y Leu Met Lys Lys Ala
 20 - # 25 - # 30

- - His Glu Ile Ser Val Leu Cys Asp Ala Glu Va - #l Ala Leu Val Val Phe
 35 - # 40 - # 45

- - Ser His Lys Gly Lys Leu Phe Glu Tyr Pro Th - #r Asp Ser Cys Met Glu
 50 - # 55 - # 60

- - Glu Ile Leu Glu Arg Tyr Glu Arg Tyr Ser Ty - #r Ala Glu Arg Gln Leu
 65 - # 70 - # 75 - # 80

- - Ile Ala Pro Glu Ser Asp Ser Asn Thr Asn Tr - #p Ser Met Glu Tyr Asn
 85 - # 90 - # 95

- - Arg Leu Lys Ala Lys Ile Glu Leu Leu Glu Ar - #g Asn Gln Arg His Tyr
 100 - # 105 - # 110

- - Leu Gly Glu Asp Leu Gln Ala Met Ser Pro Ly - #s Glu Leu Gln Asn Leu
 115 - # 120 - # 125

- - Glu Gln Gln Leu Asp Thr Ala Leu Lys His Il - #e Arg Ser Arg Lys Asn
 130 - # 135 - # 140

- - Gln Leu Met Tyr Asp Ser Ile Asn Glu Leu Gl - #n Arg Lys Glu Lys Ala
 145 1 - #50 1 - #55 1 -

#60

- - Ile Gln Glu Gln Asn Ser Met Leu Ser Lys Gl - #n Ile Lys Glu Arg

Glu 165 - # 170 - # 175

- - Asn Val Leu Arg Ala Gln Gln Glu Gln Trp As - #p Glu Gln Asn His Gly
 180 - # 185 - # 190

- - His Asn Met Pro Pro Pro Pro Pro Gln Gl - #n His Gln Ile Gln His
 195 - # 200 - # 205
 - - Pro Tyr Met Leu Ser His Gln Pro Ser Pro Ph - #e Leu Asn Met Gly Gly
 210 - # 215 - # 220
 - - Leu Tyr Gln Glu Glu Asp Gln Met Ala Met Ar - #g Arg Asn Asp Leu Asp
 225 2 - #30 2 - #35 2 -
 #40
 - - Leu Ser Leu Glu Pro Val Tyr Asn Cys Asn Le - #u Gly Arg Arg Cys
 245 - # 250 - # 255
 - - - (2) INFORMATION FOR SEQ ID NO:7:
 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1345 base - #pairs
 (B) TYPE: nucleic acid (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: cDNA
 - - (ix) FEATURE: (A) NAME/KEY: CDS
 (B) LOCATION: 149..968 - - (ix) FEATURE:
 (A) NAME/KEY: misc.sub.-- - #feature
 (B) LOCATION: 1..1345
 (D) OTHER INFORMATION: - #/note= "product = Zea mays AP1."
 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
 - - GCACGAGTCC TCCTCCTCCT CGCATCCAC CCCACCCAC CTCTCCTTA AA -
 #GCTACCTG 60
 - - CCTACCCGGC GGTTCGCGC CGCAATCGAT CGACCGGAAG AGAAAGAGCA GC -
 #TAGCTAGC 120
 - - TAGCAGATCG GAGCAGGCA ACAAGGCG ATG GGG CGC GGC AAG - #GTA CAG CTG
 172
 - # Met Gly Arg - #Gly Lys Val Gln Leu
 - # 1 - # 5
 - - AAG CGG ATA GAG AAC AAG ATA AAC CGG CAG GT - #G ACC TTC TCC AAG CGC
 220
 Lys Arg Ile Glu Asn Lys Ile Asn Arg Gln Va - #l Thr Phe Ser Lys Arg
 10 - # 15 - # 20
 - - CGG AAC GGC CTG CTC AAG AAG GCG CAC GAG AT - #C TCC GTC CTC TGC GAT
 268
 Arg Asn Gly Leu Leu Lys Lys Ala His Glu Il - #e Ser Val Leu Cys Asp
 25 - # 30 - # 35 - # 40
 - - GCC GAG GTC GCC GTC ATC GTC TTC TCC CCC AA - #G GGC AAG CTC TAC GAG
 316
 Ala Glu Val Ala Val Ile Val Phe Ser Pro Ly - #s Gly Lys Leu Tyr Glu
 45 - # 50 - # 55
 - - TAC GCC ACC GAC TCC CGC ATG GAC AAA ATT CT - #T GAA CGC TAT GAG CGA
 364
 Tyr Ala Thr Asp Ser Arg Met Asp Lys Ile Le - #u Glu Arg Tyr Glu Arg
 60 - # 65 - # 70
 - - TAT TCC TAT GCT GAA AAG GCT CTT ATT TCA GC - #T GAA TCT GAA AGT GAG
 412
 Tyr Ser Tyr Ala Glu Lys Ala Leu Ile Ser Al - #a Glu Ser Glu Ser Glu
 75 - # 80 - # 85
 - - GGA AAT TGG TGC CAC GAA TAC AGG AAA CTG AA - #G GCC AAA ATT GAG ACC
 460
 Gly Asn Trp Cys His Glu Tyr Arg Lys Leu Ly - #s Ala Lys Ile Glu Thr

90 - # 95 - # 100
 - - ATA CAA AAA TGC CAC AAG CAC CTG ATG GGA GA - #G GAT CTA GAG TCT TTG
 508
 Ile Gln Lys Cys His Lys His Leu Met Gly Gl - #u Asp Leu Glu Ser Leu
 5 105 1 - #10 1 - #15 1 -
 #20
 - - AAT CCC AAA GAG CTC CAG CAA CTA GAG CAG CA - #G CTG GAT AGC TCA
 CTG 556
 Asn Pro Lys Glu Leu Gln Gln Leu Glu Gln Gl - #n Leu Asp Ser Ser Leu
 10 125 - # 130 - # 135
 - - AAG CAC ATC AGA TCA AGG AAG AGC CAC CTT AT - #G GCC GAG TCT ATT TCT
 604
 Lys His Ile Arg Ser Arg Lys Ser His Leu Me - #t Ala Glu Ser Ile Ser
 140 - # 145 - # 150
 15 - - GAG CTA CAG AAG AAG GAG AGG TCA CTG CAG GA - #G GAG AAC AAG GCT CTG
 652
 Glu Leu Gln Lys Lys Glu Arg Ser Leu Gln Gl - #u Glu Asn Lys Ala Leu
 155 - # 160 - # 165
 - - CAG AAG GAA CTT GCG GAG AGG CAG AAG GCC GT - #C GCG AGC CCG CAG CAG
 700
 Gln Lys Glu Leu Ala Glu Arg Gln Lys Ala Va - #l Ala Ser Arg Gln Gln
 170 - # 175 - # 180
 - - CAG CAA CAG CAG CAG GTG CAG TGG GAC CAG CA - #G ACA CAT GCC CAG GCC
 748
 20 Gln Gln Gln Gln Gln Val Gln Trp Asp Gln Gl - #n Thr His Ala Gln Ala
 185 1 - #90 1 - #95 2 -
 #00
 - - CAG ACA AGC TCA TCA TCG TCC TCC TTC ATG AT - #G AGG CAG GAT CAG
 CAG 796
 30 Gln Thr Ser Ser Ser Ser Ser Ser Phe Met Me - #t Arg Gln Asp Gln Gln
 205 - # 210 - # 215
 - - GGA CTG CCG CCT CCA CAC AAC ATC TGC TTC CC - #G CCG TTG ACA ATG GGA
 844
 Gly Leu Pro Pro Pro His Asn Ile Cys Phe Pr - #o Pro Leu Thr Met Gly
 35 220 - # 225 - # 230
 - - GAT AGA GGT GAA GAG CTG GCT GCG GCG GCG GC - #G GCG CAG CAG CAG CAG
 892
 Asp Arg Gly Glu Glu Leu Ala Ala Ala Ala Al - #a Ala Gln Gln Gln Gln
 235 - # 240 - # 245
 40 - - CCA CTG CCG GGG CAG GCG CAA CCG CAG CTC CG - #C ATC GCA GGT CTG CCA
 940
 Pro Leu Pro Gly Gln Ala Gln Pro Gln Leu Ar - #g Ile Ala Gly Leu Pro
 250 - # 255 - # 260
 - - CCA TGG ATG CTG AGC CAC CTC AAT GCA T AAGG - #AGAGGG TCGATGAACA
 45 988 Pro Trp Met Leu Ser His Leu Asn Ala
 265 2 - #70
 - - CATCGACCTC CTCTCTCTCT CTCTCTCGTC ATGGATCATG ACGTACGCGT AC -
 #CATATGGT 1048
 - - TGCTGTGCCT GCCCCATCG ATCGCGAGCA ATGGCACGCT CATGCAAGTG AT -
 50 #CATTGCTC 1108
 - - CCCGTTGGTT AAACCCTAGC CTATGTTTAT GCGTCAGCA ACTAAGCTAA AC -

#TATTGTTA 1168
 - - TGTTCGCAAG AAAGGGTAAA CCCGCTAGCT GTGTAATCTT GTCCAGCTAT CA -
 #GTATGCTT 1228
 - - GTTACTGCCC AGTTACCCTT GAATCTAGCG GCGCTTTTGG TGAGAGGGTG CA -
 #GTTTACTT 1288
 - - TAAACATGGT TCGTGACTTG CTGTAAATAG TAGTATTAAT CGATTGGGC AT - #CTAAA
 1345 - - - - (2) INFORMATION FOR SEQ ID NO:8:
 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 273 amino - #acids (B) TYPE: amino acid
 (D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: protein
 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
 - - Met Gly Arg Gly Lys Val Gln Leu Lys Arg Il - #e Glu Asn Lys Ile Asn
 1 5 - # 10 - # 15
 - - Arg Gln Val Thr Phe Ser Lys Arg Arg Asn Gl - #y Leu Leu Lys Lys Ala
 20 - # 25 - # 30
 - - His Glu Ile Ser Val Leu Cys Asp Ala Glu Va - #l Ala Val Ile Val Phe
 35 - # 40 - # 45
 - - Ser Pro Lys Gly Lys Leu Tyr Glu Tyr Ala Th - #r Asp Ser Arg Met Asp
 50 - # 55 - # 60
 - - Lys Ile Leu Glu Arg Tyr Glu Arg Tyr Ser Ty - #r Ala Glu Lys Ala Leu
 65 - # 70 - # 75 - # 80
 - - Ile Ser Ala Glu Ser Glu Ser Glu Gly Asn Tr - #p Cys His Glu Tyr Arg
 85 - # 90 - # 95
 - - Lys Leu Lys Ala Lys Ile Glu Thr Ile Gln Ly - #s Cys His Lys His Leu
 100 - # 105 - # 110
 - - Met Gly Glu Asp Leu Glu Ser Leu Asn Pro Ly - #s Glu Leu Gln Gln Leu
 115 - # 120 - # 125
 - - Glu Gln Gln Leu Asp Ser Ser Leu Lys His Il - #e Arg Ser Arg Lys Ser
 130 - # 135 - # 140
 - - His Leu Met Ala Glu Ser Ile Ser Glu Leu Gl - #n Lys Lys Glu Arg Ser
 145 1 - #50 1 - #55 1 -
 #60
 - - Leu Gln Glu Glu Asn Lys Ala Leu Gln Lys Gl - #u Leu Ala Glu Arg
 Gln 165 - # 170 - # 175
 - - Lys Ala Val Ala Ser Arg Gln Gln Gln Gln Gl - #n Gln Gln Val Gln Trp
 180 - # 185 - # 190
 - - Asp Gln Gln Thr His Ala Gln Ala Gln Thr Se - #r Ser Ser Ser Ser Ser
 195 - # 200 - # 205
 - - Phe Met Met Arg Gln Asp Gln Gln Gly Leu Pr - #o Pro Pro His Asn Ile
 210 - # 215 - # 220
 - - Cys Phe Pro Pro Leu Thr Met Gly Asp Arg Gl - #y Glu Glu Leu Ala Ala
 225 2 - #30 2 - #35 2 -
 #40
 - - Ala Ala Ala Ala Gln Gln Gln Gln Pro Leu Pr - #o Gly Gln Ala Gln
 Pro 245 - # 250 - # 255
 - - Gln Leu Arg Ile Ala Gly Leu Pro Pro Trp Me - #t Leu Ser His Leu Asn
 260 - # 265 - # 270 - - Ala
 - - - - (2) INFORMATION FOR SEQ ID NO:9:
 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 779 base - #pairs

5

(B) TYPE: nucleic acid (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: cDNA
 - - (ix) FEATURE: (A) NAME/KEY: CDS
 (B) LOCATION: 10..775 - - (ix) FEATURE:
 (A) NAME/KEY: unsure (B) LOCATION: 778..779
 (D) OTHER INFORMATION: - #/note= "N = one or more
 nucleotides. - #" - - (ix) FEATURE:
 (A) NAME/KEY: misc.sub.-- - #feature
 (B) LOCATION: 1..779
 (D) OTHER INFORMATION: - #/note= "product = Arabidopsis
 thaliana - #CAL."

10

- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
 - - TTAAGAGAA ATG GGA AGG GGT AGG GTT GAA TTG AAG - # AGG ATA GAG AAC
 48

15

Met Gly Arg Gly Arg - #Val Glu Leu Lys Arg Ile Glu Asn
 1 - # 5 - # 10
 - - AAG ATC AAT AGA CAA GTG ACA TTC TCG AAA AG - #A AGA ACT GGT CTT TTG
 96
 Lys Ile Asn Arg Gln Val Thr Phe Ser Lys Ar - #g Arg Thr Gly Leu Leu
 15 - # 20 - # 25
 - - AAG AAA GCT CAG GAG ATC TCT GTT CTT TGT GA - #T GCC GAG GTT TCC CTT
 144

20

Lys Lys Ala Gln Glu Ile Ser Val Leu Cys As - #p Ala Glu Val Ser Leu
 30 - # 35 - # 40 - # 45
 - - ATT GTC TTC TCC CAT AAG GGC AAA TTG TTC GA - #G TAC TCC TCT GAA TCT
 192
 Ile Val Phe Ser His Lys Gly Lys Leu Phe Gl - #u Tyr Ser Ser Glu Ser
 50 - # 55 - # 60
 - - TGC ATG GAG AAG GTA CTA GAA CGC TAC GAG AG - #G TAT TCT TAC GCC GAG
 240

25

Cys Met Glu Lys Val Leu Glu Arg Tyr Glu Ar - #g Tyr Ser Tyr Ala Glu
 65 - # 70 - # 75
 - - AGA CAG CTG ATT GCA CCT GAC TCT CAC GTT AA - #T GCA CAG ACG AAC TGG
 288

35

Arg Gln Leu Ile Ala Pro Asp Ser His Val As - #n Ala Gln Thr Asn Trp
 80 - # 85 - # 90
 - - TCA ATG GAG TAT AGC AGG CTT AAG GCC AAG AT - #T GAG CTT TTG GAG AGA
 336

40

Ser Met Glu Tyr Ser Arg Leu Lys Ala Lys Il - #e Glu Leu Leu Glu Arg
 95 - # 100 - # 105
 - - AAC CAA AGG CAT TAT CTG GGA GAA GAG TTG GA - #A CCA ATG AGC CTC AAG
 384

45

Asn Gln Arg His Tyr Leu Gly Glu Glu Leu Gl - #u Pro Met Ser Leu Lys
 110 1 - #15 1 - #20 1 -
 #25
 - - GAT CTC CAA AAT CTG GAG CAG CAG CTT GAG AC - #T GCT CTT AAG CAC
 ATT 432

50

Asp Leu Gln Asn Leu Glu Gln Gln Leu Glu Th - #r Ala Leu Lys His Ile
 130 - # 135 - # 140
 - - CGC TCC AGA AAA AAT CAA CTC ATG AAT GAG TC - #C CTC AAC CAC CTC CAA
 480

Arg Ser Arg Lys Asn Gln Leu Met Asn Glu Se - #r Leu Asn His Leu Gln
 145 - # 150 - # 155
 - - AGA AAG GAG AAG GAG ATA CAG GAG GAA AAC AG - #C ATG CTT ACC AAA CAG
 528
 5 Arg Lys Glu Lys Glu Ile Gln Glu Glu Asn Se - #r Met Leu Thr Lys Gln
 160 - # 165 - # 170
 - - ATA AAG GAG AGG GAA AAC ATC CTA AAG ACA AA - #A CAA ACC CAA TGT GAG
 576
 10 Ile Lys Glu Arg Glu Asn Ile Leu Lys Thr Ly - #s Gln Thr Gln Cys Glu
 175 - # 180 - # 185
 - - CAG CTG AAC CGC AGC GTC GAC GAT GTA CCA CA - #G CCA CAA CCA TTT CAA
 624
 Gln Leu Asn Arg Ser Val Asp Asp Val Pro Gl - #n Pro Gln Pro Phe Gln
 190 1 - #95 2 - #00 2 -
 15 #05
 - - CAC CCC CAT CTT TAC ATG ATC GCT CAT CAG AC - #T TCT CCT TTC CTA
 AAT 672
 His Pro His Leu Tyr Met Ile Ala His Gln Th - #r Ser Pro Phe Leu Asn
 210 - # 215 - # 220
 20 - - ATG GGT GGT TTG TAC CAA GGA GAA GAC CAA AC - #G GCG ATG AGG AGG AAC
 720
 Met Gly Gly Leu Tyr Gln Gly Glu Asp Gln Th - #r Ala Met Arg Arg Asn
 225 - # 230 - # 235
 - - AAT CTG GAT CTG ACT CTT GAA CCC ATT TAC AA - #T TAC CTT GGC TGT TAC
 768
 25 Asn Leu Asp Leu Thr Leu Glu Pro Ile Tyr As - #n Tyr Leu Gly Cys Tyr
 240 - # 245 - # 250
 - - GCC GCT T GANN - # - # 779
 Ala Ala 255 - - - - (2) INFORMATION FOR SEQ ID NO:10:
 30 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 255 amino - #acids (B) TYPE: amino acid
 (D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: protein
 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
 35 - - Met Gly Arg Gly Arg Val Glu Leu Lys Arg Il - #e Glu Asn Lys Ile Asn
 1 5 - # 10 - # 15
 - - Arg Gln Val Thr Phe Ser Lys Arg Arg Thr Gl - #y Leu Leu Lys Lys Ala
 20 - # 25 - # 30
 - - Gln Glu Ile Ser Val Leu Cys Asp Ala Glu Va - #l Ser Leu Ile Val Phe
 35 - # 40 - # 45
 40 - - Ser His Lys Gly Lys Leu Phe Glu Tyr Ser Se - #r Glu Ser Cys Met Glu
 50 - # 55 - # 60
 - - Lys Val Leu Glu Arg Tyr Glu Arg Tyr Ser Ty - #r Ala Glu Arg Gln Leu
 65 - # 70 - # 75 - # 80
 - - Ile Ala Pro Asp Ser His Val Asn Ala Gln Th - #r Asn Trp Ser Met Glu
 85 - # 90 - # 95
 45 - - Tyr Ser Arg Leu Lys Ala Lys Ile Glu Leu Le - #u Glu Arg Asn Gln Arg
 100 - # 105 - # 110
 - - His Tyr Leu Gly Glu Glu Leu Glu Pro Met Se - #r Leu Lys Asp Leu Gln
 115 - # 120 - # 125
 50 - - Asn Leu Glu Gln Gln Leu Glu Thr Ala Leu Ly - #s His Ile Arg Ser Arg
 130 - # 135 - # 140

- - Lys Asn Gln Leu Met Asn Glu Ser Leu Asn Hi - #s Leu Gln Arg Lys Glu
 145 1 - #50 1 - #55 1 -
 #60
 - - Lys Glu Ile Gln Glu Glu Asn Ser Met Leu Th - #r Lys Gln Ile Lys
 5 Glu 165 - # 170 - # 175
 - - Arg Glu Asn Ile Leu Lys Thr Lys Gln Thr Gl - #n Cys Glu Gln Leu Asn
 180 - # 185 - # 190
 - - Arg Ser Val Asp Asp Val Pro Gln Pro Gln Pr - #o Phe Gln His Pro His
 195 - # 200 - # 205
 10 - - Leu Tyr Met Ile Ala His Gln Thr Ser Pro Ph - #e Leu Asn Met Gly Gly
 210 - # 215 - # 220
 - - Leu Tyr Gln Gly Glu Asp Gln Thr Ala Met Ar - #g Arg Asn Asn Leu Asp
 225 2 - #30 2 - #35 2 -
 #40
 15 - - Leu Thr Leu Glu Pro Ile Tyr Asn Tyr Leu Gl - #y Cys Tyr Ala Ala
 245 - # 250 - # 255
 - - - - (2) INFORMATION FOR SEQ ID NO:11:
 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 756 base - #pairs
 (B) TYPE: nucleic acid (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: cDNA
 - - (ix) FEATURE: (A) NAME/KEY: CDS
 (B) LOCATION: 1..754 - - (ix) FEATURE:
 (A) NAME/KEY: misc.sub.-- - #feature
 (B) LOCATION: 1..756
 (D) OTHER INFORMATION: - #/note= "product = Brassica oleracea
 CAL." - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
 - - ATG GGA AGG GGT AGG GTT GAA ATG AAG AGG AT - #A GAG AAC AAG ATC AAC
 48
 30 Met Gly Arg Gly Arg Val Glu Met Lys Arg Il - #e Glu Asn Lys Ile Asn
 1 5 - # 10 - # 15
 - - CGA CAA GTG ACG TTT TCG AAA AGA AGA GCT GG - #T CTT TTG AAG AAA GCC
 96
 Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gl - #y Leu Leu Lys Lys Ala
 35 20 - # 25 - # 30
 - - CAT GAG ATC TCG ATC CTT TGT GAT GCT GAG GT - #T TCC CTT ATT GTC TTC
 144
 His Glu Ile Ser Ile Leu Cys Asp Ala Glu Va - #l Ser Leu Ile Val Phe
 35 - # 40 - # 45
 40 - - TCC CAT AAG GGG AAA CTG TTC GAG TAC TCG TC - #T GAA TCT TGC ATG GAG
 192
 Ser His Lys Gly Lys Leu Phe Glu Tyr Ser Se - #r Glu Ser Cys Met Glu
 50 - # 55 - # 60
 - - AAG GTA CTA GAA CAC TAC GAG AGG TAC TCT TA - #C GCC GAG AAA CAG CTA
 240
 45 Lys Val Leu Glu His Tyr Glu Arg Tyr Ser Ty - #r Ala Glu Lys Gln Leu
 65 - # 70 - # 75 - # 80
 - - AAA GTT CCA GAC TCT CAC GTC AAT GCA CAA AC - #G AAC TGG TCA GTG GAA
 288
 50 Lys Val Pro Asp Ser His Val Asn Ala Gln Th - #r Asn Trp Ser Val Glu
 85 - # 90 - # 95

- - TAT AGC AGG CTT AAG GCT AAG ATT GAG CTT TT - #G GAG AGA AAC CAA AGG
336

Tyr Ser Arg Leu Lys Ala Lys Ile Glu Leu Le - #u Glu Arg Asn Gln Arg
100 - # 105 - # 110

- - CAT TAT CTG GGC GAA GAT TTA GAA TCA ATC AG - #C ATA AAG GAG CTA CAG
384

His Tyr Leu Gly Glu Asp Leu Glu Ser Ile Se - #r Ile Lys Glu Leu Gln
115 - # 120 - # 125

- - AAT CTG GAG CAG CAG CTT GAC ACT TCT CTT AA - #A CAT ATT CGC TCG AGA
432

Asn Leu Glu Gln Gln Leu Asp Thr Ser Leu Ly - #s His Ile Arg Ser Arg
130 - # 135 - # 140

- - AAA AAT CAA CTA ATG CAC GAG TCC CTC AAC CA - #C CTC CAA AGA AAG GAG
480

Lys Asn Gln Leu Met His Glu Ser Leu Asn Hi - #s Leu Gln Arg Lys Glu
145 1 - #50 1 - #55 1 -

#60

- - AAA GAA ATA CTG GAG GAA AAC AGC ATG CTT GC - #C AAA CAG ATA AGG
GAG 528

Lys Glu Ile Leu Glu Glu Asn Ser Met Leu Al - #a Lys Gln Ile Arg Glu
165 - # 170 - # 175

- - AGG GAG AGT ATC CTA AGG ACA CAT CAA AAC CA - #A TCA GAG CAG CAA AAC
576

Arg Glu Ser Ile Leu Arg Thr His Gln Asn Gl - #n Ser Glu Gln Gln Asn
180 - # 185 - # 190

- - CGC AGC CAC CAT GTA GCT CCT CAG CCG CAA CC - #G CAG TTA AAT CCT TAC
624

Arg Ser His His Val Ala Pro Gln Pro Gln Pr - #o Gln Leu Asn Pro Tyr
195 - # 200 - # 205

- - ATG GCA TCA TCT CCT TTC CTA AAT ATG GGT GG - #C ATG TAC CAA GGA GAA
672

Met Ala Ser Ser Pro Phe Leu Asn Met Gly Gl - #y Met Tyr Gln Gly Glu
210 - # 215 - # 220

- - TAT CCA ACG GCG GTG AGG AGG AAC CGT CTC GA - #T CTG ACT CTT GAA CCC
720

Tyr Pro Thr Ala Val Arg Arg Asn Arg Leu As - #p Leu Thr Leu Glu Pro
225 2 - #30 2 - #35 2 -

#40 - - ATT TAC AAC TGC AAC CTT GGT TAC TTT GCC GC - #A T GA

- # 756 Ile Tyr Asn Cys Asn Leu Gly Tyr Phe Ala Al - #a
245 - # 250

- - - (2) INFORMATION FOR SEQ ID NO:12:

- - (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 251 amino - #acids (B) TYPE: amino acid

(D) TOPOLOGY: linear

- - (ii) MOLECULE TYPE: protein

- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

- - Met Gly Arg Gly Arg Val Glu Met Lys Arg Il - #e Glu Asn Lys Ile Asn
1 5 - # 10 - # 15

- - Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gl - #y Leu Leu Lys Lys Ala
20 - # 25 - # 30

```

- - His Glu Ile Ser Ile Leu Cys Asp Ala Glu Val - #1 Ser Leu Ile Val Phe
      35      - #      40      - #      45
- - Ser His Lys Gly Lys Leu Phe Glu Tyr Ser Se - #r Glu Ser Cys Met Glu
      50      - #      55      - #      60
- - Lys Val Leu Glu His Tyr Glu Arg Tyr Ser Ty - #r Ala Glu Lys Gln Leu
      65      - # 70      - # 75      - # 80
- - Lys Val Pro Asp Ser His Val Asn Ala Gln Th - #r Asn Trp Ser Val Glu
      85 - #      90 - #      95
- - Tyr Ser Arg Leu Lys Ala Lys Ile Glu Leu Le - #u Glu Arg Asn Gln Arg
      100      - #      105      - #      110
- - His Tyr Leu Gly Glu Asp Leu Glu Ser Ile Se - #r Ile Lys Glu Leu Gln
      115      - #      120      - #      125
- - Asn Leu Glu Gln Gln Leu Asp Thr Ser Leu Ly - #s His Ile Arg Ser Arg
      130      - # 135      - # 140
- - Lys Asn Gln Leu Met His Glu Ser Leu Asn Hi - #s Leu Gln Arg Lys Glu
      145      1 - #50      1 - #55      1 -
#60
- - Lys Glu Ile Leu Glu Glu Asn Ser Met Leu Al - #a Lys Gln Ile Arg
Glu      165 - #      170 - #      175
- - Arg Glu Ser Ile Leu Arg Thr His Gln Asn Gl - #n Ser Glu Gln Gln Asn
      180      - #      185      - #      190
- - Arg Ser His His Val Ala Pro Gln Pro Gln Pr - #o Gln Leu Asn Pro Tyr
      195      - #      200      - #      205
- - Met Ala Ser Ser Pro Phe Leu Asn Met Gly Gl - #y Met Tyr Gln Gly Glu
      210      - # 215      - # 220
- - Tyr Pro Thr Ala Val Arg Arg Asn Arg Leu As - #p Leu Thr Leu Glu Pro
      225      2 - #30      2 - #35      2 -
#40 - - Ile Tyr Asn Cys Asn Leu Gly Tyr Phe Ala Al - #a
      245 - #      250
- - - - (2) INFORMATION FOR SEQ ID NO:13:
- - (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 756 base - #pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: double
      (D) TOPOLOGY: linear
- - (ii) MOLECULE TYPE: cDNA
- - (ix) FEATURE:
      (A) NAME/KEY: CDS
      (B) LOCATION: 1..451
      (ix) FEATURE:
      (A) NAME/KEY: misc.sub.-- - #feature
      (B) LOCATION: 1..756
      (D) OTHER INFORMATION: - #/note= "product = Brassica oleracea
var. botr - #ytis CAL."
- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:
- - ATG GGA AGG GGT AGG GTT GAA ATG AAG AGG AT - #A GAG AAC AAG ATC AAC
      48
Met Gly Arg Gly Arg Val Glu Met Lys Arg Il - #e Glu Asn Lys Ile Asn
      1      5 - #      10 - #      15
- - AGA CAA GTG ACG TTT TCG AAA AGA AGA GCT GG - #T CTT TTG AAG AAA GCC
      96
Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gl - #y Leu Leu Lys Lys Ala
      20      - #      25      - #      30
- - CAT GAG ATC TCG ATT CTT TGT GAT GCT GAG GT - #T TCC CTT ATT GTC TTC
      144

```

His Glu Ile Ser Ile Leu Cys Asp Ala Glu Val - #1 Ser Leu Ile Val Phe
 35 - # 40 - # 45
 - - TCC CAT AAG GGG AAA CTG TTC GAG TAC TCG TC - #T GAA TCT TGC ATG GAG
 192
 5 Ser His Lys Gly Lys Leu Phe Glu Tyr Ser Se - #r Glu Ser Cys Met Glu
 50 - # 55 - # 60
 - - AAG GTA CTA GAA CGC TAC GAG AGG TAC TCT TA - #C GCC GAG AAA CAG CTA
 240
 10 Lys Val Leu Glu Arg Tyr Glu Arg Tyr Ser Ty - #r Ala Glu Lys Gln Leu
 65 - # 70 - # 75 - # 80
 - - AAA GCT CCA GAC TCT CAC GTC AAT GCA CAA AC - #G AAC TGG TCA ATG GAA
 288
 Lys Ala Pro Asp Ser His Val Asn Ala Gln Th - #r Asn Trp Ser Met Glu
 85 - # 90 - # 95
 15 - - TAT AGC AGG CTT AAG GCT AAG ATT GAG CTT TG - #G GAG AGG AAC CAA AGG
 336
 Tyr Ser Arg Leu Lys Ala Lys Ile Glu Leu Tr - #p Glu Arg Asn Gln Arg
 100 - # 105 - # 110
 - - CAT TAT CTG GGA GAA GAT TTA GAA TCA ATC AG - #C ATA AAG GAG CTA CAG
 384
 20 His Tyr Leu Gly Glu Asp Leu Glu Ser Ile Se - #r Ile Lys Glu Leu Gln
 115 - # 120 - # 125
 - - AAT CTG GAG CAG CAG CTT GAC ACT TCT CTT AA - #A CAT ATT CGC TCC AGA
 432
 25 Asn Leu Glu Gln Gln Leu Asp Thr Ser Leu Ly - #s His Ile Arg Ser Arg
 130 - # 135 - # 140
 - - AAA AAT CAA CTA ATG CAC T AGTCCCTCAA CCACCTCCAA - #AGAAAGGAGA
 481 Lys Asn Gln Leu Met His 145 1 - #50
 - - AAGAAATACT GGAGGAAAAC AGCATGCTTG CCAAACAGAT AAAGGAGAGG GA -
 30 #GAGTATCC 541
 - - TAAGGACACA TCAAAACCAA TCAGAGCAGC AAAACCGCAG CCACCATGTA GC -
 #TCCTCAGC 601
 - - CGCAACCGCA GTTAAATCCT TACATGGCAT CATCTCCTTT CCTAAATATG GG -
 #TGGCATGT 661
 35 - - ACCAAGGAGA ATATCCAACG GCGGTGAGGA GGAACCGTCT CGATCTGACT CT -
 #TGAACCCA 721 - - TTTACAAC TG CAACCTTGGT TACTTTGCCG CATGA - #
 - # 756 - - - - (2) INFORMATION FOR SEQ ID NO:14:
 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 150 amino - #acids
 40 (B) TYPE: amino acid (D) TOPOLOGY: linear
 - - (ii) MOLECULE TYPE: protein
 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
 - - Met Gly Arg Gly Arg Val Glu Met Lys Arg Il - #e Glu Asn Lys Ile
 Asn 1 5 - # 10 - # 15
 45 - - Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gl - #y Leu Leu Lys Lys Ala
 20 - # 25 - # 30
 - - His Glu Ile Ser Ile Leu Cys Asp Ala Glu Val - #1 Ser Leu Ile Val Phe
 35 - # 40 - # 45
 - - Ser His Lys Gly Lys Leu Phe Glu Tyr Ser Se - #r Glu Ser Cys Met Glu
 50 50 - # 55 - # 60
 - - Lys Val Leu Glu Arg Tyr Glu Arg Tyr Ser Ty - #r Ala Glu Lys Gln Leu


```

65          - # 70          - # 75          - # 80
- - Lys Ala Pro Asp Ser His Val Asn Ala Gln Th - #r Asn Trp Ser Met Glu
          85 - #          90 - #          95
5  - - Tyr Ser Arg Leu Lys Ala Lys Ile Glu Leu Tr - #p Glu Arg Asn Gln Arg
          100 - #          105 - #          110
- - His Tyr Leu Gly Glu Asp Leu Glu Ser Ile Se - #r Ile Lys Glu Leu Gln
          115 - #          120 - #          125
- - Asn Leu Glu Gln Gln Leu Asp Thr Ser Leu Ly - #s His Ile Arg Ser Arg
          130 - #          135 - #          140
10 - - Lys Asn Gln Leu Met His          145          1 - #50
- - - - (2) INFORMATION FOR SEQ ID NO:15:
- - (i) SEQUENCE CHARACTERISTICS:
- - (A) LENGTH: 1500 base - #pairs
- - (B) TYPE: nucleic acid          (C) STRANDEDNESS: double
15 - - (D) TOPOLOGY: linear          - - (ii) MOLECULE TYPE: cDNA
- - (ix) FEATURE:          (A) NAME/KEY: CDS
- - (B) LOCATION: 72..1343          - - (ix) FEATURE:
- - (A) NAME/KEY: misc.sub.-- - #feature
- - (B) LOCATION: 1..1500
20 - - (D) OTHER INFORMATION: - #/note= "product = Arabidopsis
          thaliana - #LEAFY (LFY).".
- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
- - AAAGCAATCT GCTCAAAAGA GTAAAGAAAG AGAGAAAAAG AGAGTGATAG AG -
#AGAGAGAG          60
25 - - AAAAATAGAT T ATG GAT CCT GAA GGT TTC ACG AGT - #GGC TTA TTC CGG
TGG          110
          Met Asp Pro - #Glu Gly Phe Thr Ser Gly Leu Phe Arg Trp
          1 - #          5 - #          10
- - AAC CCA ACG AGA GCA TTG GTT CAA GCA CCA CC - #T CCG GTT CCA CCT CCG
30 158
Asn Pro Thr Arg Ala Leu Val Gln Ala Pro Pr - #o Pro Val Pro Pro Pro
          15 - #          20 - #          25
- - CTG CAG CAA CAG CCG GTG ACA CCG CAG ACG GC - #T GCT TTT GGG ATG CGA
206
35 Leu Gln Gln Gln Pro Val Thr Pro Gln Thr Al - #a Ala Phe Gly Met Arg
          30 - # 35          - # 40          - # 45
- - CTT GGT GGT TTA GAG GGA CTA TTC GGT CCA TA - #C GGT ATA CGT TTC TAC
254
Leu Gly Gly Leu Glu Gly Leu Phe Gly Pro Ty - #r Gly Ile Arg Phe Tyr
40          50 - #          55 - #          60
- - ACG GCG GCG AAG ATA GCG GAG TTA GGT TTT AC - #G GCG AGC ACG CTT GTG
302
Thr Ala Ala Lys Ile Ala Glu Leu Gly Phe Th - #r Ala Ser Thr Leu Val
          65 - #          70 - #          75
45 - - GGT ATG AAG GAC GAG GAG CTT GAA GAG ATG AT - #G AAT AGT CTC TCT CAT
350
Gly Met Lys Asp Glu Glu Leu Glu Glu Met Me - #t Asn Ser Leu Ser His
          80 - #          85 - #          90
- - ATC TTT CGT TGG GAG CTT CTT GTT GGT GAA CG - #G TAC GGT ATC AAA GCT
50 398
Ile Phe Arg Trp Glu Leu Leu Val Gly Glu Ar - #g Tyr Gly Ile Lys Ala

```

95 - # 100 - # 105
 - - GCC GTT AGA GCT GAA CGG AGA CGA TTG CAA GA - #A GAG GAG GAA GAG GAA
 446
 Ala Val Arg Ala Glu Arg Arg Arg Leu Gln Gl - #u Glu Glu Glu Glu Glu
 5 110 1 - #15 1 - #20 1 -
 #25
 - - TCT TCT AGA CGC CGT CAT TTG CTA CTC TCC GC - #C GCT GGT GAT TCC
 GGT 494
 Ser Ser Arg Arg Arg His Leu Leu Leu Ser Al - #a Ala Gly Asp Ser Gly
 10 130 - # 135 - # 140
 - - ACT CAT CAC GCT CTT GAT GCT CTC TCC CAA GA - #A GAT GAT TGG ACA GGG
 542
 Thr His His Ala Leu Asp Ala Leu Ser Gln Gl - #u Asp Asp Trp Thr Gly
 145 - # 150 - # 155
 15 - - TTA TCT GAG GAA CCG GTG CAG CAA CAA GAC CA - #G ACT GAT GCG GCG GGG
 590
 Leu Ser Glu Glu Pro Val Gln Gln Gln Asp Gl - #n Thr Asp Ala Ala Gly
 160 - # 165 - # 170
 - - AAT AAC GGC GGA GGA GGA AGT GGT TAC TGG GA - #C GCA GGT CAA GGA AAG
 638
 Asn Asn Gly Gly Gly Gly Ser Gly Tyr Trp As - #p Ala Gly Gln Gly Lys
 175 - # 180 - # 185
 - - ATG AAG AAG CAA CAG CAG CAG AGA CGG AGA AA - #G AAA CCA ATG CTG ACG
 686
 25 Met Lys Lys Gln Gln Gln Gln Arg Arg Arg Ly - #s Lys Pro Met Leu Thr
 190 1 - #95 2 - #00 2 -
 #05
 - - TCA GTG GAA ACC GAC GAA GAC GTC AAC GAA GG - #T GAG GAT GAC GAC
 GGG 734
 Ser Val Glu Thr Asp Glu Asp Val Asn Glu Gl - #y Glu Asp Asp Asp Gly
 210 - # 215 - # 220
 - - ATG GAT AAC GGC AAC GGA GGT AGT GGT TTG GG - #G ACA GAG AGA CAG AGG
 782
 Met Asp Asn Gly Asn Gly Gly Ser Gly Leu Gl - #y Thr Glu Arg Gln Arg
 225 - # 230 - # 235
 - - GAG CAT CCG TTT ATC GTA ACG GAG CCT GGG GA - #A GTG GCA CGT GGC AAA
 830
 Glu His Pro Phe Ile Val Thr Glu Pro Gly Gl - #u Val Ala Arg Gly Lys
 240 - # 245 - # 250
 40 - - AAG AAC GGC TTA GAT TAT CTG TTC CAC TTG TA - #C GAA CAA TGC CGT GAG
 878
 Lys Asn Gly Leu Asp Tyr Leu Phe His Leu Ty - #r Glu Gln Cys Arg Glu
 255 - # 260 - # 265
 - - TTC CTT CTT CAG GTC CAG ACA ATT GCT AAA GA - #C CGT GGC GAA AAA TGC
 926
 45 Phe Leu Leu Gln Val Gln Thr Ile Ala Lys As - #p Arg Gly Glu Lys Cys
 270 2 - #75 2 - #80 2 -
 #85
 - - CCC ACC AAG GTG ACG AAC CAA GTA TTC AGG TA - #C GCG AAG AAA TCA
 50 GGA 974
 Pro Thr Lys Val Thr Asn Gln Val Phe Arg Ty - #r Ala Lys Lys Ser Gly

20
25
30
35
40
45
50

```

                290 - #                295 - #                300
- - GCG AGT TAC ATA AAC AAG CCT AAA ATG CGA CA - #C TAC GTT CAC TGT TAC
1022
Ala Ser Tyr Ile Asn Lys Pro Lys Met Arg Hi - #s Tyr Val His Cys Tyr
                305 - #                310 - #                315
- - GCT CTC CAC TGC CTA GAC GAA GAA GCT TCA AA - #T GCT CTC AGA AGA GCG
1070
Ala Leu His Cys Leu Asp Glu Glu Ala Ser As - #n Ala Leu Arg Arg Ala
                320 - #                325 - #                330
- - TTT AAA GAA CGC GGT GAG AAC GTT GGC TCA TG - #G CGT CAG GCT TGT TAC
1118
Phe Lys Glu Arg Gly Glu Asn Val Gly Ser Tr - #p Arg Gln Ala Cys Tyr
                335 - #                340 - #                345
- - AAG CCA CTT GTG AAC ATC GCT TGT CGT CAT GG - #C TGG GAT ATA GAC GCC
1166
Lys Pro Leu Val Asn Ile Ala Cys Arg His Gl - #y Trp Asp Ile Asp Ala
350                3 - #55                3 - #60                3 -
#65
- - GTC TTT AAC GCT CAT CCT CGT CTC TCT ATT TG - #G TAT GTT CCA ACA
AAG 1214
Val Phe Asn Ala His Pro Arg Leu Ser Ile Tr - #p Tyr Val Pro Thr Lys
                370 - #                375 - #                380
- - CTG CGT CAG CTT TGC CAT TTG GAG CGG AAC AA - #T GCG GTT GCT GCG GCT
1262
Leu Arg Gln Leu Cys His Leu Glu Arg Asn As - #n Ala Val Ala Ala Ala
                385 - #                390 - #                395
- - GCG GCT TTA GTT GGC GGT ATT AGC TGT ACC GG - #A TCG TCG ACG TCT GGA
1310
Ala Ala Leu Val Gly Gly Ile Ser Cys Thr Gl - #y Ser Ser Thr Ser Gly
                400 - #                405 - #                410
- - CGT GGT GGA TGC GGC GGC GAC GAC TTG CGT TT - #C TAGTTTGGTT TGGGTAGTT
G 1363 Arg Gly Gly Cys Gly Gly Asp Asp Leu Arg Ph - #e
                415 - #                420
- - TGGTTTGT TT AGTCGTTATC CTAATTA ACT ATTAGTCTTT AATTAGTCT TC -
#TTGGCTAA 1423
- - TTTATTTTTC TTTTTTGTG AAAACCTTTA ATTTGTTATG GCTAATTTGT TA -
#TACACGCA 1483
- - GTTTTCTTAA TGCGTTA - # - #
- # 1500 - - - - (2) INFORMATION FOR SEQ ID NO:16:
- - (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 424 amino - #acids (B) TYPE: amino acid
      (D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: protein
- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
- - Met Asp Pro Glu Gly Phe Thr Ser Gly Leu Ph - #e Arg Trp Asn Pro Thr
1                5 - #                10 - #                15
- - Arg Ala Leu Val Gln Ala Pro Pro Pro Val Pr - #o Pro Pro Leu Gln Gln
20 - #                25 - #                30
- - Gln Pro Val Thr Pro Gln Thr Ala Ala Phe Gl - #y Met Arg Leu Gly Gly
35 - #                40 - #                45
- - Leu Glu Gly Leu Phe Gly Pro Tyr Gly Ile Ar - #g Phe Tyr Thr Ala Ala
```

50 - # 55 - # 60
- - Lys Ile Ala Glu Leu Gly Phe Thr Ala Ser Th - #r Leu Val Gly Met Lys
65 - # 70 - # 75 - # 80
- - Asp Glu Glu Leu Glu Glu Met Met Asn Ser Le - #u Ser His Ile Phe Arg
5 85 - # 90 - # 95
- - Trp Glu Leu Leu Val Gly Glu Arg Tyr Gly Il - #e Lys Ala Ala Val Arg
100 - # 105 - # 110
- - Ala Glu Arg Arg Arg Leu Gln Glu Glu Glu Gl - #u Glu Glu Ser Ser Arg
115 - # 120 - # 125
10 - - Arg Arg His Leu Leu Leu Ser Ala Ala Gly As - #p Ser Gly Thr His His
130 - # 135 - # 140
- - Ala Leu Asp Ala Leu Ser Gln Glu Asp Asp Tr - #p Thr Gly Leu Ser Glu
145 1 - #50 1 - #55 1 -
#60
15 - - Glu Pro Val Gln Gln Gln Asp Gln Thr Asp Al - #a Ala Gly Asn Asn
Gly 165 - # 170 - # 175
- - Gly Gly Gly Ser Gly Tyr Trp Asp Ala Gly Gl - #n Gly Lys Met Lys Lys
180 - # 185 - # 190
- - Gln Gln Gln Gln Arg Arg Arg Lys Lys Pro Me - #t Leu Thr Ser Val Glu
195 - # 200 - # 205
- - Thr Asp Glu Asp Val Asn Glu Gly Glu Asp As - #p Asp Gly Met Asp Asn
210 - # 215 - # 220
- - Gly Asn Gly Gly Ser Gly Leu Gly Thr Glu Ar - #g Gln Arg Glu His Pro
225 2 - #30 2 - #35 2 -
#40
- - Phe Ile Val Thr Glu Pro Gly Glu Val Ala Ar - #g Gly Lys Lys Asn
Gly 245 - # 250 - # 255
- - Leu Asp Tyr Leu Phe His Leu Tyr Glu Gln Cy - #s Arg Glu Phe Leu Leu
260 - # 265 - # 270
30 - - Gln Val Gln Thr Ile Ala Lys Asp Arg Gly Gl - #u Lys Cys Pro Thr Lys
275 - # 280 - # 285
- - Val Thr Asn Gln Val Phe Arg Tyr Ala Lys Ly - #s Ser Gly Ala Ser Tyr
290 - # 295 - # 300
- - Ile Asn Lys Pro Lys Met Arg His Tyr Val Hi - #s Cys Tyr Ala Leu His
305 3 - #10 3 - #15 3 -
#20
- - Cys Leu Asp Glu Glu Ala Ser Asn Ala Leu Ar - #g Arg Ala Phe Lys
Glu 325 - # 330 - # 335
- - Arg Gly Glu Asn Val Gly Ser Trp Arg Gln Al - #a Cys Tyr Lys Pro Leu
340 - # 345 - # 350
- - Val Asn Ile Ala Cys Arg His Gly Trp Asp Il - #e Asp Ala Val Phe Asn
355 - # 360 - # 365
- - Ala His Pro Arg Leu Ser Ile Trp Tyr Val Pr - #o Thr Lys Leu Arg Gln
370 - # 375 - # 380
45 - - Leu Cys His Leu Glu Arg Asn Asn Ala Val Al - #a Ala Ala Ala Ala Leu
385 3 - #90 3 - #95 4 -
#00
- - Val Gly Gly Ile Ser Cys Thr Gly Ser Ser Th - #r Ser Gly Arg Gly
Gly 405 - # 410 - # 415
50 - - Cys Gly Gly Asp Asp Leu Arg Phe 420
- - - - (2) INFORMATION FOR SEQ ID NO:17:

```

- - (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 1656 base - #pairs
      (B) TYPE: nucleic acid (C) STRANDEDNESS: double
      (D) TOPOLOGY: linear - - (ix) FEATURE:
      (A) NAME/KEY: CDS (B) LOCATION: 1..1651
- - (ix) FEATURE: (A) NAME/KEY: misc.sub.-- - #feature
      (B) LOCATION: 1..1656
      (D) OTHER INFORMATION: - #/note= "domain = ecdysone receptor
          ligand bi - #nding domain."
- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
- - ATG CCG CCG GAA TGC GTC GTC CCG GAG AAC CA - #A TGT GCG ATG AAG CCG
    48
Met Arg Pro Glu Cys Val Val Pro Glu Asn Gl - #n Cys Ala Met Lys Arg
  1      5 - #      10 - #      15
- - CGC GAA AAG AAG GCC CAG AAG GAG AAG GAC AA - #A ATG ACC ACT TCG CCG
    96
Arg Glu Lys Lys Ala Gln Lys Glu Lys Asp Ly - #s Met Thr Thr Ser Pro
      20 - #      25 - #      30
- - AGC TCT CAG CAT GGC GGC AAT GGC AGC TTG GC - #C TCT GGT GGC GGC CAA
    144
Ser Ser Gln His Gly Gly Asn Gly Ser Leu Al - #a Ser Gly Gly Gly Gln
      35 - #      40 - #      45
- - GAC TTT GTT AAG AAG GAG ATT CTT GAC CTT AT - #G ACA TGC GAG CCG CCC
    192
Asp Phe Val Lys Lys Glu Ile Leu Asp Leu Me - #t Thr Cys Glu Pro Pro
      50 - #      55 - #      60
- - CAG CAT GCC ACT ATT CCG CTA CTA CCT GAT GA - #A ATA TTG GCC AAG TGT
    240
Gln His Ala Thr Ile Pro Leu Leu Pro Asp Gl - #u Ile Leu Ala Lys Cys
      65 - # 70 - # 75 - # 80
- - CAA GCG CGC AAT ATA CCT TCC TTA ACG TAC AA - #T CAG TTG GCC GTT ATA
    288
Gln Ala Arg Asn Ile Pro Ser Leu Thr Tyr As - #n Gln Leu Ala Val Ile
      85 - #      90 - #      95
- - TAC AAG TTA ATT TGG TAC CAG GAT GGC TAT GA - #G CAG CCA TCT GAA GAG
    336
Tyr Lys Leu Ile Trp Tyr Gln Asp Gly Tyr Gl - #u Gln Pro Ser Glu Glu
      100 - #      105 - #      110
- - GAT CTC AGG CGT ATA ATG AGT CAA CCC GAT GA - #G AAC GAG AGC CAA ACG
    384
Asp Leu Arg Arg Ile Met Ser Gln Pro Asp Gl - #u Asn Glu Ser Gln Thr
      115 - #      120 - #      125
- - GAC GTC AGC TTT CGG CAT ATA ACC GAG ATA AC - #C ATA CTC ACG GTC CAG
    432
Asp Val Ser Phe Arg His Ile Thr Glu Ile Th - #r Ile Leu Thr Val Gln
      130 - #      135 - #      140
- - TTG ATT GTT GAG TTT GCT AAA GGT CTA CCA GC - #G TTT ACA AAG ATA CCC
    480
Leu Ile Val Glu Phe Ala Lys Gly Leu Pro Al - #a Phe Thr Lys Ile Pro
      145 1 - #50 1 - #55 1 -
#60

```

- - CAG GAG GAC CAG ATC ACG TTA CTA AAG GCC TG - #C TCG TCG GAG GTG
 ATG 528
 Gln Glu Asp Gln Ile Thr Leu Leu Lys Ala Cy - #s Ser Ser Glu Val Met
 165 - # 170 - # 175
 - - ATG CTG CGT ATG GCA CGA CGC TAT GAC CAC AG - #C TCG GAC TCA ATA TTC
 576
 Met Leu Arg Met Ala Arg Arg Tyr Asp His Se - #r Ser Asp Ser Ile Phe
 180 - # 185 - # 190
 - - TTC GCG AAT AAT AGA TCA TAT ACG CGG GAT TC - #T TAC AAA ATG GCC GGA
 624
 Phe Ala Asn Asn Arg Ser Tyr Thr Arg Asp Se - #r Tyr Lys Met Ala Gly
 195 - # 200 - # 205
 - - ATG GCT GAT AAC ATT GAA GAC CTG CTG CAT TT - #C TGC CGC CAA ATG TTC
 672
 Met Ala Asp Asn Ile Glu Asp Leu Leu His Ph - #e Cys Arg Gln Met Phe
 210 - # 215 - # 220
 - - TCG ATG AAG GTG GAC AAC GTC GAA TAC GCG CT - #T CTC ACT GCC ATT GTG
 720
 Ser Met Lys Val Asp Asn Val Glu Tyr Ala Le - #u Leu Thr Ala Ile Val
 225 2 - #30 2 - #35 2 -
 #40
 - - ATC TTC TCG GAC CGG CCG GGC CTG GAG AAG GC - #C CAA CTA GTC GAA
 GCG 768
 Ile Phe Ser Asp Arg Pro Gly Leu Glu Lys Al - #a Gln Leu Val Glu Ala
 245 - # 250 - # 255
 - - ATC CAG AGC TAC TAC ATC GAC ACG CTA CGC AT - #T TAT ATA CTC AAC CGC
 816
 Ile Gln Ser Tyr Tyr Ile Asp Thr Leu Arg Il - #e Tyr Ile Leu Asn Arg
 260 - # 265 - # 270
 - - CAC TGC GGC GAC TCA ATG AGC CTC GTC TTC TA - #C GCA AAG CTG CTC TCG
 864
 His Cys Gly Asp Ser Met Ser Leu Val Phe Ty - #r Ala Lys Leu Leu Ser
 275 - # 280 - # 285
 - - ATC CTC ACC GAG CTG CGT ACG CTG GGC AAC CA - #G AAC GCC GAG ATG TGT
 912
 Ile Leu Thr Glu Leu Arg Thr Leu Gly Asn Gl - #n Asn Ala Glu Met Cys
 290 - # 295 - # 300
 - - TTC TCA CTA AAG CTC AAA AAC CGC AAA CTG CC - #C AAG TTC CTC GAG GAG
 960
 Phe Ser Leu Lys Leu Lys Asn Arg Lys Leu Pr - #o Lys Phe Leu Glu Glu
 305 3 - #10 3 - #15 3 -
 #20
 - - ATC TGG GAC GTT CAT GCC ATC CCG CCA TCG GT - #C CAG TCG CAC CTT
 CAG 1008
 Ile Trp Asp Val His Ala Ile Pro Pro Ser Va - #l Gln Ser His Leu Gln
 325 - # 330 - # 335
 - - ATT ACC CAG GAG GAG AAC GAG CGT CTC GAG CG - #G GCT GAG CGT ATG CGG
 1056
 Ile Thr Gln Glu Glu Asn Glu Arg Leu Glu Ar - #g Ala Glu Arg Met Arg
 340 - # 345 - # 350
 - - GCA TCG GTT GGG GGC GCC ATT ACC GCC GGC AT - #T GAT TGC GAC TCT GCC

1104
Ala Ser Val Gly Gly Ala Ile Thr Ala Gly Il - #e Asp Cys Asp Ser Ala
355 - # 360 - # 365
- - TCC ACT TCG GCG GCG GCA GCC GCG GCC CAG CA - #T CAG CCT CAG CCT CAG
1152
Ser Thr Ser Ala Ala Ala Ala Ala Ala Gln Hi - #s Gln Pro Gln Pro Gln
370 - # 375 - # 380
- - CCC CAG CCC CAA CCC TCC TCC CTG ACC CAG AA - #C GAT TCC CAG CAC CAG
1200
Pro Gln Pro Gln Pro Ser Ser Leu Thr Gln As - #n Asp Ser Gln His Gln
385 3 - #90 3 - #95 4 -
#00
- - ACA CAG CCG CAG CTA CAA CCT CAG CTA CCA CC - #T CAG CTG CAA GGT
CAA 1248
Thr Gln Pro Gln Leu Gln Pro Gln Leu Pro Pr - #o Gln Leu Gln Gly Gln
405 - # 410 - # 415
- - CTG CAA CCC CAG CTC CAA CCA CAG CTT CAG AC - #G CAA CTC CAG CCA CAG
1296
Leu Gln Pro Gln Leu Gln Pro Gln Leu Gln Th - #r Gln Leu Gln Pro Gln
420 - # 425 - # 430
- - ATT CAA CCA CAG CCA CAG CTC CTT CCC GTC TC - #C GCT CCC GTG CCC GCC
1344
Ile Gln Pro Gln Pro Gln Leu Leu Pro Val Se - #r Ala Pro Val Pro Ala
435 - # 440 - # 445
- - TCC GTA ACC GCA CCT GGT TCC TTG TCC GCG GT - #C AGT ACG AGC AGC GAA
1392
Ser Val Thr Ala Pro Gly Ser Leu Ser Ala Va - #l Ser Thr Ser Ser Glu
450 - # 455 - # 460
- - TAC ATG GGC GGA AGT GCG GCC ATA GGA CCC AT - #C ACG CCG GCA ACC ACC
1440
Tyr Met Gly Gly Ser Ala Ala Ile Gly Pro Il - #e Thr Pro Ala Thr Thr
465 4 - #70 4 - #75 4 -
#80
- - AGC AGT ATC ACG GCT GCC GTT ACC GCT AGC TC - #C ACC ACA TCA GCG
GTA 1488
Ser Ser Ile Thr Ala Ala Val Thr Ala Ser Se - #r Thr Thr Ser Ala Val
485 - # 490 - # 495
- - CCG ATG GGC AAC GGA GTT GGA GTC GGT GTT GG - #G GTG GGC GGC AAC GTC
1536
Pro Met Gly Asn Gly Val Gly Val Gly Val Gl - #y Val Gly Gly Asn Val
500 - # 505 - # 510
- - AGC ATG TAT GCG AAC GCC CAG ACG GCG ATG GC - #C TTG ATG GGT GTA GCC
1584
Ser Met Tyr Ala Asn Ala Gln Thr Ala Met Al - #a Leu Met Gly Val Ala
515 - # 520 - # 525
- - CTG CAT TCG CAC CAA GAG CAG CTT ATC GGG GG - #A GTG GCG GTT AAG TCG
1632
Leu His Ser His Gln Glu Gln Leu Ile Gly Gl - #y Val Ala Val Lys Ser
530 - # 535 - # 540
- - GAG CAC TCG ACG ACT GCA T AGCAG - # - #
1656 Glu His Ser Thr Thr Ala 545 5 - #50

- - - (2) INFORMATION FOR SEQ ID NO:18:
 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 550 amino - #acids (B) TYPE: amino acid
 (D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: protein
 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

- - Met Arg Pro Glu Cys Val Val Pro Glu Asn Gl - #n Cys Ala Met Lys Arg
 1 5 - # 10 - # 15
 - - Arg Glu Lys Lys Ala Gln Lys Glu Lys Asp Ly - #s Met Thr Thr Ser Pro
 20 - # 25 - # 30
 - - Ser Ser Gln His Gly Gly Asn Gly Ser Leu Al - #a Ser Gly Gly Gly Gln
 35 - # 40 - # 45

- - Asp Phe Val Lys Lys Glu Ile Leu Asp Leu Me - #t Thr Cys Glu Pro Pro
 50 - # 55 - # 60
 - - Gln His Ala Thr Ile Pro Leu Leu Pro Asp Gl - #u Ile Leu Ala Lys Cys
 65 - # 70 - # 75 - # 80
 - - Gln Ala Arg Asn Ile Pro Ser Leu Thr Tyr As - #n Gln Leu Ala Val Ile
 85 - # 90 - # 95

- - Tyr Lys Leu Ile Trp Tyr Gln Asp Gly Tyr Gl - #u Gln Pro Ser Glu Glu
 100 - # 105 - # 110
 - - Asp Leu Arg Arg Ile Met Ser Gln Pro Asp Gl - #u Asn Glu Ser Gln Thr
 115 - # 120 - # 125
 - - Asp Val Ser Phe Arg His Ile Thr Glu Ile Th - #r Ile Leu Thr Val Gln
 130 - # 135 - # 140

- - Leu Ile Val Glu Phe Ala Lys Gly Leu Pro Al - #a Phe Thr Lys Ile Pro
 145 1 - #50 1 - #55 1 -
 #60

- - Gln Glu Asp Gln Ile Thr Leu Leu Lys Ala Cy - #s Ser Ser Glu Val
 Met 165 - # 170 - # 175

- - Met Leu Arg Met Ala Arg Arg Tyr Asp His Se - #r Ser Asp Ser Ile Phe
 180 - # 185 - # 190
 - - Phe Ala Asn Asn Arg Ser Tyr Thr Arg Asp Se - #r Tyr Lys Met Ala Gly
 195 - # 200 - # 205

- - Met Ala Asp Asn Ile Glu Asp Leu Leu His Ph - #e Cys Arg Gln Met Phe
 210 - # 215 - # 220
 - - Ser Met Lys Val Asp Asn Val Glu Tyr Ala Le - #u Leu Thr Ala Ile Val
 225 2 - #30 2 - #35 2 -

#40
 - - Ile Phe Ser Asp Arg Pro Gly Leu Glu Lys Al - #a Gln Leu Val Glu
 Ala 245 - # 250 - # 255

- - Ile Gln Ser Tyr Tyr Ile Asp Thr Leu Arg Il - #e Tyr Ile Leu Asn Arg
 260 - # 265 - # 270
 - - His Cys Gly Asp Ser Met Ser Leu Val Phe Ty - #r Ala Lys Leu Leu Ser
 275 - # 280 - # 285

- - Ile Leu Thr Glu Leu Arg Thr Leu Gly Asn Gl - #n Asn Ala Glu Met Cys
 290 - # 295 - # 300
 - - Phe Ser Leu Lys Leu Lys Asn Arg Lys Leu Pr - #o Lys Phe Leu Glu Glu
 305 3 - #10 3 - #15 3 -

#20
 - - Ile Trp Asp Val His Ala Ile Pro Pro Ser Va - #l Gln Ser His Leu
 Gln 325 - # 330 - # 335

- - Ile Thr Gln Glu Glu Asn Glu Arg Leu Glu Ar - #g Ala Glu Arg Met Arg
 340 - # 345 - # 350
 - - Ala Ser Val Gly Gly Ala Ile Thr Ala Gly Il - #e Asp Cys Asp Ser Ala
 355 - # 360 - # 365
 - - Ser Thr Ser Ala Ala Ala Ala Ala Gln Hi - #s Gln Pro Gln Pro Gln
 370 - # 375 - # 380
 - - Pro Gln Pro Gln Pro Ser Ser Leu Thr Gln As - #n Asp Ser Gln His Gln
 385 3 - #90 3 - #95 4 -
 #00
 - - Thr Gln Pro Gln Leu Gln Pro Gln Leu Pro Pr - #o Gln Leu Gln Gly
 Gln 405 - # 410 - # 415
 - - Leu Gln Pro Gln Leu Gln Pro Gln Leu Gln Th - #r Gln Leu Gln Pro Gln
 420 - # 425 - # 430
 - - Ile Gln Pro Gln Pro Gln Leu Leu Pro Val Se - #r Ala Pro Val Pro Ala
 435 - # 440 - # 445
 - - Ser Val Thr Ala Pro Gly Ser Leu Ser Ala Va - #l Ser Thr Ser Ser Glu
 450 - # 455 - # 460
 - - Tyr Met Gly Gly Ser Ala Ala Ile Gly Pro Il - #e Thr Pro Ala Thr Thr
 465 4 - #70 4 - #75 4 -
 #80
 - - Ser Ser Ile Thr Ala Ala Val Thr Ala Ser Se - #r Thr Thr Ser Ala
 Val 485 - # 490 - # 495
 - - Pro Met Gly Asn Gly Val Gly Val Gly Val Gl - #y Val Gly Gly Asn Val
 500 - # 505 - # 510
 - - Ser Met Tyr Ala Asn Ala Gln Thr Ala Met Al - #a Leu Met Gly Val Ala
 515 - # 520 - # 525
 - - Leu His Ser His Gln Glu Gln Leu Ile Gly Gl - #y Val Ala Val Lys Ser
 530 - # 535 - # 540
 - - Glu His Ser Thr Thr Ala 545 5 - #50
 - - - - (2) INFORMATION FOR SEQ ID NO:19:
 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 855 base - #pairs
 (B) TYPE: nucleic acid (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear - - (ix) FEATURE:
 (A) NAME/KEY: CDS (B) LOCATION: 1..853
 - - (ix) FEATURE: (A) NAME/KEY: misc.sub.-- - #feature
 (B) LOCATION: 1..855
 (D) OTHER INFORMATION: - #/note= "domain = glucocorticoid
 receptor - #ligand binding domain."
 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
 - - ACA AAG AAA AAA ATC AAA GGG ATT CAG CAA GC - #C ACT GCA GGA GTC TCA
 48
 Thr Lys Lys Lys Ile Lys Gly Ile Gln Gln Al - #a Thr Ala Gly Val Ser
 1 5 - # 10 - # 15
 - - CAA GAC ACT TCG GAA AAT CCT AAC AAA ACA AT - #A GTT CCT GCA GCA TTA
 96
 Gln Asp Thr Ser Glu Asn Pro Asn Lys Thr Il - #e Val Pro Ala Ala Leu
 20 - # 25 - # 30
 - - CCA CAG CTC ACC CCT ACC TTG GTG TCA CTG CT - #G GAG GTG ATT GAA CCC
 144
 Pro Gln Leu Thr Pro Thr Leu Val Ser Leu Le - #u Glu Val Ile Glu Pro

35 - # 40 - # 45
 - - GAG GTG TTG TAT GCA GGA TAT GAT AGC TCT GT - #T CCA GAT TCA GCA TGG
 192
 Glu Val Leu Tyr Ala Gly Tyr Asp Ser Ser Va - #1 Pro Asp Ser Ala Trp
 5 50 - # 55 - # 60
 - - AGA ATT ATG ACC ACA CTC AAC ATG TTA GGT GG - #G CGT CAA GTG ATT GCA
 240
 Arg Ile Met Thr Thr Leu Asn Met Leu Gly Gl - #y Arg Gln Val Ile Ala
 65 - # 70 - # 75 - # 80
 10 - - GCA GTG AAA TGG GCA AAG GCG ATA CTA GGC TT - #G AGA AAC TTA CAC CTC
 288
 Ala Val Lys Trp Ala Lys Ala Ile Leu Gly Le - #u Arg Asn Leu His Leu
 85 - # 90 - # 95
 - - GAT GAC CAA ATG ACC CTG CTA CAG TAC TCA TG - #G ATG TTT CTC ATG GCA
 15 336
 Asp Asp Gln Met Thr Leu Leu Gln Tyr Ser Tr - #p Met Phe Leu Met Ala
 100 - # 105 - # 110
 - - TTT GCC TTG GGT TGG AGA TCA TAC AGA CAA TC - #A AGC GGA AAC CTG CTC
 384
 Phe Ala Leu Gly Trp Arg Ser Tyr Arg Gln Se - #r Ser Gly Asn Leu Leu
 115 - # 120 - # 125
 - - TGC TTT GCT CCT GAT CTG ATT ATT AAT GAG CA - #G AGA ATG TCT CTA CCC
 432
 Cys Phe Ala Pro Asp Leu Ile Ile Asn Glu Gl - #n Arg Met Ser Leu Pro
 130 - # 135 - # 140
 25 - - TGC ATG TAT GAC CAA TGT AAA CAC ATG CTG TT - #T GTC TCC TCT GAA TTA
 480
 Cys Met Tyr Asp Gln Cys Lys His Met Leu Ph - #e Val Ser Ser Glu Leu
 145 1 - #50 1 - #55 1 -
 #60
 - - CAA AGA TTG CAG GTA TCC TAT GAA GAG TAT CT - #C TGT ATG AAA ACC
 TTA 528
 Gln Arg Leu Gln Val Ser Tyr Glu Glu Tyr Le - #u Cys Met Lys Thr Leu
 165 - # 170 - # 175
 35 - - CTG CTT CTC TCC TCA GTT GCT AAG GAA GGT CT - #G AAG AGC CAA GAG TTA
 576
 Leu Leu Leu Ser Ser Val Ala Lys Glu Gly Le - #u Lys Ser Gln Glu Leu
 180 - # 185 - # 190
 - - TTT GAT GAG ATT CGA ATG ACT TAT ATC AAA GA - #G CTA GGA AAA GCC ATC
 40 624
 Phe Asp Glu Ile Arg Met Thr Tyr Ile Lys Gl - #u Leu Gly Lys Ala Ile
 195 - # 200 - # 205
 - - GTC AAA AGG GAA GGG AAC TCC AGT CAG AAC TG - #G CAA CGG TTT TAC CAA
 672
 45 Val Lys Arg Glu Gly Asn Ser Ser Gln Asn Tr - #p Gln Arg Phe Tyr Gln
 210 - # 215 - # 220
 - - CTG ACA AAG CTT CTG GAC TCC ATG CAT GAG GT - #G GTT GAG AAT CTC CTT
 720
 Leu Thr Lys Leu Leu Asp Ser Met His Glu Va - #1 Val Glu Asn Leu Leu
 50 225 2 - #30 2 - #35 2 -
 #40

- - ACC TAC TGC TTC CAG ACA TTT TTG GAT AAG AC - #C ATG AGT ATT GAA
TTC 768
Thr Tyr Cys Phe Gln Thr Phe Leu Asp Lys Th - #r Met Ser Ile Glu Phe
245 - # 250 - # 255
- - CCA GAG ATG TTA GCT GAA ATC ATC ACT AAT CA - #G ATA CCA AAA TAT TCA
816
Pro Glu Met Leu Ala Glu Ile Ile Thr Asn Gl - #n Ile Pro Lys Tyr Ser
260 - # 265 - # 270
- - AAT GGA AAT ATC AAA AAG CTT CTG TTT CAT CA - #A AAA T GA
- # 855 Asn Gly Asn Ile Lys Lys Leu Leu Phe His Gl - #n Lys
275 - # 280
- - - (2) INFORMATION FOR SEQ ID NO:20:
- - (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 284 amino - #acids (B) TYPE: amino acid
(D) TOPOLOGY: linear - - (ii) MOLECULE TYPE: protein
- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
- - Thr Lys Lys Lys Ile Lys Gly Ile Gln Gln Al - #a Thr Ala Gly Val Ser
1 5 - # 10 - # 15
- - Gln Asp Thr Ser Glu Asn Pro Asn Lys Thr Il - #e Val Pro Ala Ala Leu
20 20 - # 25 - # 30
- - Pro Gln Leu Thr Pro Thr Leu Val Ser Leu Le - #u Glu Val Ile Glu Pro
35 35 - # 40 - # 45
- - Glu Val Leu Tyr Ala Gly Tyr Asp Ser Ser Va - #l Pro Asp Ser Ala Trp
50 50 - # 55 - # 60
- - Arg Ile Met Thr Thr Leu Asn Met Leu Gly Gl - #y Arg Gln Val Ile Ala
65 65 - # 70 - # 75 - # 80
- - Ala Val Lys Trp Ala Lys Ala Ile Leu Gly Le - #u Arg Asn Leu His Leu
85 85 - # 90 - # 95
- - Asp Asp Gln Met Thr Leu Leu Gln Tyr Ser Tr - #p Met Phe Leu Met Ala
100 100 - # 105 - # 110
- - Phe Ala Leu Gly Trp Arg Ser Tyr Arg Gln Se - #r Ser Gly Asn Leu Leu
115 115 - # 120 - # 125
- - Cys Phe Ala Pro Asp Leu Ile Ile Asn Glu Gl - #n Arg Met Ser Leu Pro
130 130 - # 135 - # 140
- - Cys Met Tyr Asp Gln Cys Lys His Met Leu Ph - #e Val Ser Ser Glu Leu
145 145 - #50 1 - #55 1 -
#60
- - Gln Arg Leu Gln Val Ser Tyr Glu Glu Tyr Le - #u Cys Met Lys Thr
Leu 165 - # 170 - # 175
- - Leu Leu Leu Ser Ser Val Ala Lys Glu Gly Le - #u Lys Ser Gln Glu Leu
180 180 - # 185 - # 190
- - Phe Asp Glu Ile Arg Met Thr Tyr Ile Lys Gl - #u Leu Gly Lys Ala Ile
195 195 - # 200 - # 205
- - Val Lys Arg Glu Gly Asn Ser Ser Gln Asn Tr - #p Gln Arg Phe Tyr Gln
210 210 - # 215 - # 220
- - Leu Thr Lys Leu Leu Asp Ser Met His Glu Va - #l Val Glu Asn Leu Leu
225 225 - #30 2 - #35 2 -
#40
- - Thr Tyr Cys Phe Gln Thr Phe Leu Asp Lys Th - #r Met Ser Ile Glu
Phe 245 - # 250 - # 255
- - Pro Glu Met Leu Ala Glu Ile Ile Thr Asn Gl - #n Ile Pro Lys Tyr Ser

```

260      - #      265      - #      270
- - Asn Gly Asn Ile Lys Lys Leu Leu Phe His Gl - #n Lys
275      - #      280
5  - - - - (2) INFORMATION FOR SEQ ID NO:21:
- -      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 50 base - #pairs      (B) TYPE: nucleic acid
          (C) STRANDEDNESS: double      (D) TOPOLOGY: linear
- -      (ix) FEATURE:      (A) NAME/KEY: misc.sub.-- - #feature
          (B) LOCATION: 1..50
10      (D) OTHER INFORMATION: - #/note= "element = copper inducible
          regulatory - #element (ACE1 binding site)."
```

- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
- - AGCTTAGCGA TCGCTCTTTT CCGCTGAACC GTTCCAGCAA AAAAGACTAG - # 50
- - - - (2) INFORMATION FOR SEQ ID NO:22:

15 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 base - #pairs (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double (D) TOPOLOGY: linear
- - (ix) FEATURE: (A) NAME/KEY: misc.sub.-- - #feature
 (B) LOCATION: 1..19
20 (D) OTHER INFORMATION: - #/note= "element = tet operator."

- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
- - ACTCTATCAG TGATAGAGT - # - #
- - # 19 - - - - (2) INFORMATION FOR SEQ ID NO:23:

25 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 29 base - #pairs (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double (D) TOPOLOGY: linear
- - (ix) FEATURE: (A) NAME/KEY: misc.sub.-- - #feature
 (B) LOCATION: 1..29
 (D) OTHER INFORMATION: - #/note= "element = ecdysone response
 element."

30 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
- - GATCCGACAA GGGTTCAATG CACTTGTC - # - #
29 - - - - (2) INFORMATION FOR SEQ ID NO:24:

35 - - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 371 base - #pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double (D) TOPOLOGY: linear
- - (ix) FEATURE: (A) NAME/KEY: misc.sub.-- - #feature
 (B) LOCATION: 1..371
 (D) OTHER INFORMATION: - #/note= "element = heat shock
 inducible - #regulatory element (HSP81-1 promoter)."

40 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
- - GTGGAGTCTC GAAACGAAAA GAACTTTCTG GAATTCGTTT GCTCACAAAG CT -
45 #AAAAACGG 60
- - TTGATTTTCAT CGAAATACGG CGTCGTTTTC AAAGAACAAT CCAGAAATCA CT -
#GGTTTTCC 120
- - TTTATTTCAA AAGAAGAGAC TAGAACTTTA TTTCTCCTCT ATAAATCAC TT -
#TGTTTTTC 180
50 - - CCTCTCTTCT TCATAATCA ACAAACAAT CACAAATCTC TCGAAACGCT CT -
#CGAAGTTC 240

```

- - CAAATTTTCT CTTAGCATTC TCTTTCGTTT CTCGTTTGCG TTGAATCAAA GT -
#TCGTTGCG      300
- - ATGGCGGATG TTCAGATGGC TGATGCAGAG ACTTTTGCTT TCCAAGCTGA GA -
#TTAACCAG      360
5  - - CTTCTTAGCT T          - #          - #          - #          371
- - - - (2) INFORMATION FOR SEQ ID NO:25:
- - (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 29 base - #pairs          (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single              (D) TOPOLOGY: linear
10 - - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
- - GGATCCGGAT CAAAAATGGG AAGGGGTAG          - #          - #
      29 - - - - (2) INFORMATION FOR SEQ ID NO:26:
- - (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 30 base - #pairs          (B) TYPE: nucleic acid
15 - - (C) STRANDEDNESS: single              (D) TOPOLOGY: linear
- - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
- - GGATCCGCTG CGGCGAAGCA GCCAAGGTTG          - #          - #
      30

```

SEQ ID NO:27 and SEQ ID NO:28

Arabidopsis SEP1 cDNA and Arabidopsis SEP1 amino acid sequence

```

      20      40      60      80      100
ATGGGAAGAG GAAGAGTAGA GCTGAAGAGG ATAGAGAACA AAATCAACAG ACAAGTAACG TTGCAAAGC GTAGGAACGG TTGTGTGAAG AAAGCTTATG
TACCCCTTCTC CTCTCATCTC CGACTTCTCC TATCTCTTGT TTAGTTGTGC TGTTTCATGC AAACGTTTCG CATCCTTGCC AAACAACATC TTTCGAATAC
M G R G R V E L K R I E N K I N R Q V T F A K R R N G L L K K A Y>

      120      140      160      180      200
AATTGTCTGT TCTCTGTGAT GCTGAAGTTG CTCTCATCAT CTCTCCAAC CGTGGAAAGC TCTATGAGTT TTGAGCTCC TCAACATGC TCAAGACACT
TTAACAGACA AGAGACACTA CGACTTCAAC GAGAGTAGTA GAAGAGGTTG GCACCTTTCG AGATACTCAA AACGTGAGG AGTTTGATCG AGTTCTGTGA
E L S V L C D A E V A L I I F S N R G K L Y E F C S S S N M L K T L>

      220      240      260      280      300
TGATCGGTAC CAGAAATGCA GCTATGGATC CATTGAAGTC AACAACAAAC CTGCCAAAGA ACTTGAGAAC AGCTACAGAG AATATCTGAA GCTTAAGGGT
ACTAGCCATG GTCTTTACGT CGATACCTAG GTAACCTCAG TTGTTGTTTG GACGGTTTCT TGAACCTCTG TCGATGTCTC TTATAGACTT CGAATTCCCA
D R Y Q K C S Y G S I E V N N K P A K E L E N S Y R E Y L K L K G>

      320      340      360      380      400
AGATATGAGA ACCTTCAACG TCAACAGAGA AATCTTCTTG GGGAGGATTT AGGACCTTTG AATTCAAAGG AGTTAGAGCA GCTTGAGCGT CAATGGAGC
TCTATACTCT TGGAAGTTGC AGTTGTCTCT TTAGAAGAAC CCTCTCTAAA TCCTGGAAGC TTAAGTTTCC TCAATCTCGT CGAAGCTGCA GTTGACCTGC
R Y E N L Q R Q Q R N L L G E D L G P L N S K E L E Q L E R Q L D>

      420      440      460      480      500
GCTCTCTCAA GCAAGTTCGG TCCATCAAGA CACAGTACAT GCTTGACCAG CTCTCGGATC TTCAAAATAA AGAGCAAATG TTGCTTGAAA CCAATAGAGC
CGAGAGAGTT CGTTCAAGCC AGGTAGTTCT GTGTCATGTA CGAAGTGGTC GAGAGCCTAG AAGTTTATT TCTCGTTTAC AACGAACTTT GGTATCTCTG
G S L K Q V R S I K T Q Y M L D Q L S D L Q N K E Q M L L E T N R A>

      520      540      560      580      600
TTTGGCAATG AAGCTGGATG ATATGATTGG TGTGAGAAGT CATCATATGG GAGGATGGGA AGGCGGTGAA CAGAATGTTA CCTACGCGCA TCATCAAGCT
AAACCGTTAC TTCGACCTAC TATACTAACC ACACTCTTCA GTAGTATACC CTCCTACCTT TCCGCCACTT GTCTTACAAT GGATGCGCGT AGTAGTTCGA
L A M K L D D M I G V R S H H M G G W E G G E Q N V T Y A H H Q A>

      620      640      660      680      700
CAGTCTCAGG GACTATACCA GCCTCTTGAA TGCAATCCAA CTCTGCAAT GGGGTATGAT AATCCAGTAT GCTCTGAGCA AATCACTGCG ACAACACAAG
GTCAGAGTCC CTGATATGGT CGGAGAACCT ACGTTAGGTT GAGACGTTTA CCCCATACTA TTAGGTCATA CGAGACTCGT TTAGTGACGC TGTTGTGTTT
Q S Q G L Y Q P L E C N P T L Q M G Y D N P V C S E Q I T A T T Q>

      720      740
CTCAGGCGCA GCCGGGAAAC GGTACATTC CAGGATGGAT GCTCTGA
GAGTCCGCGT CGGCCCTTTG CCAATGTAAG GTCCTACCTA CGAGACT
A Q A Q P G N G Y I P G W M L *>

```

SEQ ID NO:29 and SEQ ID NO:30

Arabidopsis SEP2 cDNA and Arabidopsis SEP2 amino acid sequence

```

5      20      40      60      80      100
ATGGAAGAG GAAGAGTAGA GCTCAAGAGG ATAGAGAACA AAATCAACAG ACAAGTGACG TTGCTAAAC GTAGAAATGG TTGCTGAAA AAAGCTTATG
TACCTTCTC CTCTCATCT CGAGTTCTCC TATCTCTGT TTTAGTTGTC TGTTCACTGC AAACGATTGG CATCTTTACC AAACGACTTT TTTCGAATAC
M G R G R V E L K R I E N K I N R Q V T F A K R R N G L L K K A Y>

10     120     140     160     180     200
AGCTTCTGT TCTCTGCGAT GCTGAAGTCT CTCTCATCGT CTCTCCAAC CGTGGCAAGC TCTACGAGTT CTGCAGCACC TCCAACATGC TCAAGACACT
TCGAAAGACA AGAGACGCTA CGACTTCAGA GAGAGTAGCA GAAGAGGTTG GCACCGTTCC AGATGCTCAA GACGTCGTGG AGTTGTAGAC AGTTCTGTGA
E L S V L C D A E V S L I V F S N R G K L Y E F C S T S N M L K T L>

15     220     240     260     280     300
GGAAAGGTAT CAGAAGTGTA GCTATGGCTC CATTGAAGTC AACAAACAAC CTGCTAAAGA GCTTGAGAAC AGCTACAGAG AGTACTTGAA GCTGAAAGGT
CCTTTCCATA GTCTTCACAT CGATACCGAG GTAACCTCAG TTGTTGTTTG GACGATTCTC CGAAGCTTTG TCGATGCTCT TCATGAAGTT CGACTTTCCA
E R Y Q K C S Y G S I E V N N K P A K E L E N S Y R E Y L K L K G>

20     320     340     360     380     400
AGATATGAAA ATCTGCAACG TCAGCAGAGA AATCTTCTTG GAGAGGATCT TGGACCTCTG AATTCAAAGG AGCTAGAGCA GCTTGAGCGT CAACATAGAG
TCTATACTTT TAGACGTTGC AGTCGTCTCT TTAGAAGAAC CTCTCCTAGA ACCTCGGAGC TTAAGTTTCC TCGATCTCGT CGAAGCTGCA GTTGATCTGC
R Y E N L Q R Q Q R N L L G E D L G P L N S K E L E Q L E R Q L D>

25     420     440     460     480     500
GCTCTCTGAA GCAAGTTTCG TGCATCAAGA CACAGTATAT GCTTGACCAG CTCTCTGATC TTCAAGGTAA GGAGCATATC TTGCTTGATG CCAACAGAGC
CGAGAGACTT CGTTCAAGCG ACGTAGTTCT GTGTCTATATA CGAAGCTGGT GAGAGACTAG AAGTTCCATT CCTCGTATAG AACGAAGTAC GGTGTCTCG
G S L K Q V R C I K T Q Y M L D Q L S D L Q G K E H I L L D A N R A>

30     520     540     560     580     600
TTTGTAATG AAGCTGGAAG ATATGATCGG CGTGAGACAT CACCATATAG GAGGAGGATG GGAAGGTGGT GATCAACAGA ATATTGCTTA TGGACATCCT
AAACAGTTAC TTGACCTTTC TATACTAGCC GCACTCTGTA GTGGTATATC CTCTCCTAC CTTTCCACCA CTAGTTGTCT TATAACGGAT ACCTGTAGGA
L S M K L E D M I G V R H H H I G G G W E G G D Q Q N I A Y G H P>

35     620     640     660     680     700
CAGGCTCATT CTCAGGACT ATACCAATCT CTGGAATGTG ATCCACATTT GCAAAATTGA TATAGCCATC CAGTGTGCTC AGAGCAAATG GCTGTGACGG
GTCCGAGTAA GAGTCCCTGA TATGGTTAGA GAACCTTAC TAGGGTGAAA CGTTTAACCT ATATCGGTAG GTCACACGAG TCTCGTTTAC CGACACTGCC
Q A H S Q G L Y Q S L E C D P T L Q I G Y S H P V C S E Q M A V T>

40     720     740
TGCAAGGTCA GTCCCAACAA GGAACGGCT ACATCCCTGG CTGGATGCTG TGA
ACGTTCCAGT CAGGGTTGTT CCTTTGCCGA TGTAGGACC GACCTACGAC ACT
V Q G Q S Q Q G N G Y I P G W M L *>

```

SEQ ID NO:31 and SEQ ID NO:32

Arabidopsis SEP3 cDNA and Arabidopsis SEP3 amino acid sequence

```

45     20     40     60     80     100
ATGGAAGAG GGAGAGTAGA ATTGAAGAGG ATAGAGAACA AGATCAATAG GCAAGTGACG TTGCTAAAGA GAAGGAATGG TCTTTTGAAG AAAGCATACG
TACCTTCTC CTCTCATCT TAACCTCTCC TATCTCTGT TCTAGTTATC CGTCACTGC AAACGTTTCT CTTCCTTACC AGAAACTTC TTTCGTATGC
M G R G R V E L K R I E N K I N R Q V T F A K R R N G L L K K A Y>

50     120     140     160     180     200
AGCTTTCAGT TCTATGTGAT GCAGAAGTTG CTCTCATCAT CTCTCAAAT AGAGGAAAGC TGTACGAGTT TTGAGTAGT TCGAGCATGC TTCGGCACT
TCGAAAGTCA AGATACACTA CGTCTTCAAC GAGAGTAGTA GAAGAGTTTA TCTCCTTTCG ACATGCTCAA AACGTCATCA AGCTCGTACG AAGCCTGTGA
E L S V L C D A E V A L I I F S N R G K L Y E F C S S S S M L R T L>

55     220     240     260     280     300
GGAGAGGTAC CAAAGTGTA ACTATGGAGC ACCAGAACCC AATGTGCCTT CAAGAGAGGC CTTAGCAGTT GAACCTTAGTA GCCAGCAGGA GTATCTCAAG
CCTTCCATG GTTTTCACAT TGATACCTCG TGGTCTTGGG TTACACGGAA GTTCTCTCCG GAATCGTCAA CTTGAATCAT CGGTCTCTCT CATAGAGTTC
E R Y Q K C N Y G A P E P N V P S R E A L A V E L S S Q Q E Y L K>

60     320     340     360     380     400
CTTAAGGAGC GTTATGACGC CTTACAAAGA ACCCAAAGGA ATCTGTTGGG AGAAGATCTT GGACCTCTAA GTACAAAGGA GCTTGAGTCA CTTGAGAGAC
GAATTCCTCG CAATACGCG GAATGTTTCT TGGGTTTCCT TAGACAACCC TCTTCTAGAA CCTGAGAGAT CATGTTTCTC CGAAGTCACT GAATCTCTG
L K E R Y D A L Q R T Q R N L L G E D L G P L S T K E L E S L E R>

65     420     440     460     480     500
AGCTTGATTC TTCCTGAAG CAGATCAGAG CTCTCAGGAC ACAGTTTATG CTTGACCAGC TCAACGATCT TCAGAGTAAG TTAGCTGATG GGTATCAGAT
TCGAACTAAG AAGGAACCTC GTCTAGTCTC GAGAGTCTCG TGTCAAATAC GAATCGTGG AGTTGCTAGA AGTCTCATTC AATCGACTAC CCATAGTCTA

```

Q L D S S L K Q I R A L R T Q P M L D Q L N D L Q S K L A D G Y Q M>

520 540 560 580 600
GCCACTCCAG CTGAACCCTA ACCAAGAAGA GGTGATCAC TACGGTCGTC ATCATCATCA ACAACAACAA CACTCCCAAG CTTTCTTCCA GCCTTTGGAA
CGGTGAGGTC GACTTGGGAT TGGTTCTTCT CCAACTAGTG ATGCCAGCAG TAGTAGTAGT TGTGTGTTT GTGAGGGTTC GAAAGAAGGT CGGAAACCTT
P L Q L N P N Q E E V D H Y G R H H H Q Q Q Q H S Q A F F Q P L E>
620 640 660 680 700
TGTGAACCCA TTCTTCAGAT CGGTATCAG GGGCAGCAAG ATGGAATGGG AGCAGGACCA AGTGTGAATA ATTACATGTT GGGTGGTTA CCTTATGACA
ACACTTGGGT AAGAAGTCTA GCCCATAGTC CCCGTCGTTT TACCTTACCC TCGTCCTGTT TCACACTTAT TAATGTACAA CCCAACCAAT GGAATACTGT
C E P I L Q I G Y Q G Q Q D G M G A G P S V N N Y M L G W L P Y D>
CCAACCTCTAT TTGA
GGTTGAGATA AACT
T N S I *>

SEQ ID NO:33 and SEQ ID NO:34

Arabidopsis AGL20 cDNA and Arabidopsis AGL20 amino acid sequence

20 40 60 80 100
ATGTTGAGGG GCAAACTCA GATGAAGAGA ATAGAGAATG CAACAAGCAG ACAAGTGACT TTCTCCAAA GAAGGAATGG TTTGTTGAAG AAAGCCTTTG
TACCCTCCC CGTTTGAAGT CTACTTCTCT TATCTCTTAC GTTGTTCGTC TGTTCAGTGA AAGAGGTTTT CTTCCTTACC AAACAACCTC TTTCGGAAAC
M V R G K T Q M K R I E N A T S R Q V T F S K R R N G L L K K A F>
120 140 160 180 200
AGCTCTCAGT GCTTGTGAT GCTGAAGTTT CTCCTTTCAT CTCTCTCTCT AAAGGCAAA TTTATGAATT CGCCAGCTCC AATATGCAAG ATACCATAGA
TCGAGAGTCA CGAAACACTA CGACTTCAAA GAGAATAGTA GAAGAGAGGA TTTCGTTTG AAATACTTAA GCGGTCGAGG TTATACGTTT TATGGTATCT
E L S V L C D A E V S L I I F S P K G K L Y E F A S S N M Q D T I D>
220 240 260 280 300
TCGTTATCTG AGGCATACTA AGGATCGAGT CAGCACCAAA CCGGTTTCTG AAGAAAATAT GCAGCATTG AAATATGAAG CAGCAACAT GATGAAGAAA
AGCAATAGAC TCCGTATGAT TCCTAGCTCA GTCTGTGTTT GGCCAAAGAC TTCTTTTATA CGTCGTAAC TTTATCTTCT GTCTTTGTA CTACTTCTTT
R Y L R H T K D R V S T K P V S E E N M Q H L K Y E A A N M M K K>
320 340 360 380 400
ATTGAACAAC TCGAAGCTTC TAAACGTAAA CTCTTGGGAG AAGGCATAGG AACATGCTCA ATCGAGGAGC TGCAACAGAT TGAGCAACAG CTTGAGAAAA
TAACCTGTGG AGCTTCGAAG ATTTGCTATT GAGAACCTTC TTCCGTATCC TTGTACGAGT TAGCTCTCG ACGTTGTCTA ACTCGTTGTC GAATCTCTTT
I E Q L E A S K R K L L G E G I G T C S I E E L Q Q I E Q Q L E K>
420 440 460 480 500
GTGTCAAATG TATTGAGCA AGAAGACTC AAGTGTTTAA GGAACAAATT GAGCAGCTCA AGCAAAAGGA GAAAGCTCTA GCTGCAGAAA ACGAGAAGCT
CACAGTTTAC ATAAGCTCGT TCTTTCGAG TTCACAAATT CCTTGTTTAA CTCGTCGAGT TCGTTTTCCT CTTTCGAGAT CGACGCTTTT TGCTCTTCGA
S V K C I R A R K T Q V F K E Q I E Q L K Q K E K A L A A E N E K L>
520 540 560 580 600
CTCTGAAAAG TGGGGATCTC ATGAAAGCGA AGTTTGGTCA AATAAGAATC AAGAAGTAC TGAAGAGGT GATGAAGAGA GTAGCCCAAG TTCTGAAGTA
GAGACTTTTC ACCCTAGAG TACTTTCGCT TCAAACAGT TTATTCCTAG TTCTTTCATG ACCTTCTCCA CTACTTCTCT CATCGGGTTC AAGACTTCAT
S E K W G S H E S E V W S N K N Q E S T G R G D E E S S P S S E V>
620 640
GAGACGCAAT TGTTCATTGG GTTACCTTGT TCTTCAAGAA AGTGA
CTCTGCGTTA ACAAGTAACC CAATGGAACA AGAAGTTCTT TCACT
E T Q L F I G L P C S S R K *>

SEQ ID NO:35 and SEQ ID NO:36

Arabidopsis AGL22 cDNA and Arabidopsis AGL22 amino acid sequence

20 40 60 80 100
ATGGCGAGAG AAAAATTCA GATCAGGAAG ATCGACAACG CAACGGCGAG ACAAGTGACG TTTTCGAAAC GAAGAAGAGG GCTTTTCAAG AAAGCTGAAG
TACCGCTCTC TTTTCTAAGT CTAGTCCTTC TAGCTGTTCG GTTCCGCTC TGTTCAGTGC AAAAGCTTTG CTCTTCTCC CGAAAGTTC TTTCGACTTC
M A R E K I Q I R K I D N A T A R Q V T F S K R R R G L F K K A E>
120 140 160 180 200
AACTCTCCGT TCTCTGCGAC GCCGATGTC CTCTCATCAT CTCTCTTCC ACCGGAAAAC TGTTCGAGTT CTGTAGCTCC AGCATGAAGG AAGTCTTAGA
TTGAGAGGCA AGAGACGCTG CGGCTACAGC GAGAGTAGTA GAAGAGAAGG TGGCCTTTTG ACAAGCTCAA GACATCGAGG TCGTACTTCC TTCAGGATCT
E L S V L C D A D V A L I I F S S T G K L F E F C S S S M K E V L E>
220 240 260 280 300

GAGGCATAAC TTNCAGTCAA AGAACTTGA GAAGCTTCAT CAGCCATCTC TTGAGTTACA GCTGGTTGAG AACAGTGATC ACGCCCGAAT GAGTAAAGAA
CTCCGTATTG AANGTCAGTT TCTTGAACCT CTTGCAAGTA GTCGGTAGAG AACTCAATGT CGACCAACTC TTGCACTAG TCGGGGCTTA CTCATTCTCT
R H N X Q S K N L E K L H Q P S L E L Q L V E N S D H A R M S K E>

320 340 360 380 400
ATTGCGGACA AGAGCCACCG ACTAAGGCAA ATGAGAGGAG AGGAACCTCA AGGACTTGAC ATTGAAGAGC TTCAGCAGCT AGAGAAGGCC CTTGAAACTG
TAACGCTGT TCTCGTGGC TGATTCCGTT TACTCTCCTC TCCTTGAAGT TCCTGAACTG TAACCTCTCG AAGTCGTCGA TCTCTCCGG GAACCTTGAC
I A D K S H R L R Q M R G E E L Q G L D I E E L Q Q L E K A L E T>

420 440 460 480 500
GTTTGACCGG TGTGATTGAA ACAAGAGTG ACAAGATTAT GAGTGAGATC AGCGAACTTC AGAAAAAGGG AATGCAATTG ATGGATGAGA ACAAGCGGTT
CAAACTCGGC ACACAACTT TGTCTCTCAC TGTCTTAATA CTCACTTAG TCGCTGAAG TCTTTTCC CTTACGTTAAC TACCTACTCT TGTTCGCCAA
G L T R V I E T K S D K I M S E I S E L Q K K G M Q L M D E N K R L>

520 540 560 580 600
GAGGCAGCAA GTATGTGTCT TACCCTCTCT GTTGATAACA AATCCCTTTC TTTTGTCTAC CATTAACGTA CACACTCCTA AATTTAATCC CCAGTTGTCT
CTCCGTCGTT CATACACAGA ATGGGAGAGA CAACATTTGT TTAGGGAAG AAAACAGATG GTAATTGCAT GTGTGAGGAT TTAATTTAGG GGTCAACAGA
R Q Q V C V L P S L L I T N P F L L S T I N V H T P K F N P Q L S>

620
ACAACACATA TGTGTGATCA TACTGTGAGA TAA
TGTGTGTAT ACAAACTAGT ATGACACTCT ATT
T T H M F D H T V R *>

SEQ ID NO:37 and SEQ ID NO:38

Arabidopsis AGL24 cDNA and Arabidopsis AGL24 amino acid sequence

20 40 60 80 100
ATGGCGAGAG AGAAGATAAG GATAAAGAAG ATTGATAACA TAACAGCGAG ACAAGTTACT TTCTCAAAGA GAAGAAGAGG AATCTTCAAG AAAGCCGATG
TACCGCTCTC TCTTCTATTC CTATTTCTTC TAACTATTGT ATTGTCGCTC TGTTCATGA AAGAGTTTCT CTCTCTCTCC TTAGAAGTTC TTTCGGCTAC
M A R E K I R I K K I D N I T A R Q V T F S K R R R G I F K K A D>

120 140 160 180 200
AACTTTCAGT TCTTTGCGAT GCTGATGTTG CTCTCATCAT CTCTCTGCC ACCGGAAAGC TCTTCGAGTT CTCAGCTCA AGAATGAGAG ACATATTGGG
TTGAAAGTCA AGAAAGCGCTA CGACTACAAC GAGAGTAGTA GAAGAGACGG TGGCCTTTTCG AGAAGCTCAA GAGGTGAGT TCTTACTCTC TGTATAACCC
E L S V L C D A D V A L I I F S A T G K L F E F S S S R M R D I L G>

220 240 260 280 300
AAGGTATAGT CTTATGCAA GTAACATCAA CAAATTGATG GATCCACCTT CTACTCATCT CCGGCTTGAG AATTGTAACC TCTCCAGACT AAGTAAGGAA
TTCCATATCA GAAGTACGTT CATTGTAGTT GTTTAACTAC CTAGGTGGAA GATGAGTAGA GGCCGAACCT TTAACATTGG AGAGGTCTGA TTCATTCCCT
R Y S L H A S N I N K L M D P P S T H L R L E N C N L S R L S K E>

320 340 360 380 400
GTCGAAGACA AAACCAAGCA GCTACGGAAA CTGAGAGGAG AGGATCTTGA TGGATTGAAC TTAGAAGAGT TGCAGCGGCT GGAGAAACTA CTTGAATCCG
CAGCTTCTGT TTTGGTTGCT CGATGCCCTT GACTCTCCTC TCCTAGAAGT ACCTAACTTG AATCTTCTCA ACGTCGCCGA CCTCTTTGAT GAACCTAGGC
V E D K T K Q L R K L R G E D L D G L N L E E L Q R L E K L L E S>

420 440 460 480 500
GACTTAGCCG TGTGTCTGAA AAGAAGGCCG AGTGTGTGAT GAGCCAAATT TTCTCAGTTG AGAAACGGGG ATCGGAATTG GTGGATGAGA ATAAGAGACT
CTGAATCGGC ACACAGACTT TTCTTCCCGC TCACACACTA CTCGGTTTAA AAGAGTGAAC TCTTTGCCCC TAGCCTTAAC CACCTACTCT TATTCTCTGA
G L S R V S E K K G E C V M S Q I F S L E K R G S E L V D E N K R L>

520 540 560 580 600
GAGGGATAAA CTAGAGAGCT TGGAAAGGCC AAAACTGACG ACGCTTAAAG AGGCTTTGGA GACAGAGTCG GTGACCACAA ATGTGTCAAG CTACGACAGT
CTCCCTATTT GATCTCTGCA ACCCTTCCCG TTTTGACTGC TGCGAATTTT TCCGAAACCT CTGTCTCAGC CACTGGTGTI TACACAGTTC GATGCTGTCA
R D K L E T L E R A K L T T L K E A L E T E S V T T N V S S Y D S>

620 640 660
GGAACTCCCC TTGAGATGA CTCCGACACT TCCCTGAAGC TTGGGCTTCC ATCTTGGGAA TGA
CCTTGAGGGG AACTCTACT GAGGCTGTGA AGGGACTTCG AACCCGAAGG TAGAACCTT ACT
G T P L E D D S D T S L K L G L P S W E *>

SEQ ID NO:39 and SEQ ID NO:40

Arabidopsis AGL27 cDNA and Arabidopsis AGL27 amino acid sequence

20 40 60 80 100
ATGGGAAGAA GAAAAATCGA GATCAAGCGA ATCGAGAACAA AAAGCAGTCG ACAAGTCACT TTCTCCAAAC GACGCAATGG TCTCATCGAC AAAGCTCGAC
TACCTCTCTT CTTTITAGCT CTAGTTTCGCT TAGCTCTTGT TTTCTCAGC TGTTCAGTGA AAGAGGTTTG CTGCGTTACC AGAGTAGCTG TTTCTAGCTG

M G R R K I E I K R I E N K S S R Q V T F S K R R N G L I D K A R>

120 140 160 180 200
 AACTTTTCGAT TCTCTGTGAA TCCTCCGTCG CTGTTGTGCT CGTATCTGCC TCCGGAAAAC TCTATGACTC TTCCTCCGGT GACGACATTT CCAAGATCAT
 TTGAAAGCTA AGAGACACTT AGGAGGCAGC GACAACAGCA GCATAGACGG AGGCCTTTTG AGATACTGAG AAGGAGGCCA CTGCTGTAAA GGTTCCTAGTA
 Q L S I L C E S S V A V V V V S A S G K L Y D S S S G D D I S K I I>

220 240 260 280 300
 TGATCGTTAT GAAATACAAC ATGCTGATGA ACTTAGAGCC TTAGATCTTG AAGAAAAAAT TCAGAATTAT CTTCCACACA AGGAGTTACT AGAAACAGTC
 ACTAGCAATA CTTTATGTTG TACGACTACT TGAATCTCGG AATCTAGAAC TTCTTTTITA AGTCTTAATA GAAGGTGTGT TCCTCAATGA TCTTTGTCAG
 D R Y E I Q H A D E L R A L D L E E K I Q N Y L P H K E L L E T V>

320 340 360 380 400
 CAAAGCAAGC TTGAAGAACC AAATGTCGAT AATGTAAGTG TAGATTCTCT AATTCTCTCT GAGGAACAAC TTGAGACTGC TCTGTCCGTA AGTAGAGCTA
 GTTTCGTTTG AACTTCTTGG TTACAGCTA TTACATTAC ATCTAAGAGA TTAAGAGAC CTCTTGTGTT AACTCTGACG AGACAGGCAT TCATCTCGAT
 Q S K L E E P N V D N V S V D S L I S L E E Q L E T A L S V S R A>

420 440 460 480 500
 GGAAGGCAGA ACTGATGATG GAGTATATCG AGTCCCTTAA AGAAAAGGAG AAATTGCTGA GAGAAGAGAA CCAGGTTCTG GCTAGCCAGC TGTGAGAGAA
 CCTTCCGTCT TGACTACTAC CTCATATAGC TCAGGGAATT TCTTTTCTC TTTAACGACT CTCTTCTCTT GGTCCAAGAC CGATCGGTG ACAGTCTCTT
 R K A E L M M E Y I E S L K E K E K L L R E E N Q V L A S Q L S E K>

520 540 560 580 600
 GAAAGGTATG TCTCACCAGT GAAAGATACT CAAAACCGA TGGGAAAGAA TACGTTGCTG GCAACAGATG ATGAGAGAGG AATGTTTCCG GGAAGTAGCT
 CTTTCCATAC AGAGTGGCTA CTTTCTATGA GTTTTGGGCT ACCCTTCTT ATGCAACGAC CGTTGCTAC TACTCTCTCC TTACAAAGGC CCTTCATCGA
 K G M S H R * K I L K T R W E R I R C W Q Q M M R E E C F R E V A>

620 640 660 680
 CCGGCAACAA AATACGGAG ACTCTCCGC TGCTCAATTA GCCACCATCA TCAACGGCTG AGTTTTCACC TTAACCTCAA AGCCTGA
 GGCGTTGTT TTATGGCCTC TGAGAGGGCG ACGAGTTAAT CGGTGGTAGT AGTTGCCGAC TCAAAAGTGG AATTGAGTT TCGGACT
 P A T K Y R R L S R C S I S H H H Q R L S F H L K L K A *>

SEQ ID NO:41

Arabidopsis SEP1 genomic sequence

-2981 -2961 -2941 -2921 -2901
 CAGATCTCTT GGCATGTGTC GAAATGTGG AGATCTTAAG AATGTAGCTT GTGGCCGTTG CAAAGGAACA GGAACAATCA AATCAGGAGG ATTCTTTGGT
 GTCTAGAGAA CCGTACACAG CTTTACACAC TCTAGAAATC TTACATCGAA CACCGGCAAC GTTCTCTTGT CTTGTGTAGT TTAGCTCTCC TAAGAAACCA

-2881 -2861 -2841 -2821 -2801
 TTCAGTGACT CATCAACAC AAGATCAGTG GCTTGCGATA ATTGCCAAGC CAAAGGTTGT TTCCCTTGCC CTGAATGCTC AAAATCTTGA CCATTTTCTC
 AAGTCACTGA GTAGTTTGTG TTCTAGTCAC CGAACGCTAT TAACGGTTCG GTTTCACAA AAGGGAACGG GACTTACGAG TTTTGAAGCT GGTAAAGAG

-2781 -2761 -2741 -2721 -2701
 GGTATTTTAT AGTTGTTTCA TCTTCTTGAC ACTATGATAA GTGTAATCGG TCCATTGGTA ATGGTAATGT TAAAGTTGAA GAATGTCTTG TTTATTTCGAG
 CCATAAATA TCAACAAAGT AGAAGAACTG TGATACTATT CACATTAGCC AGGTAACCAT TACCATTACA ATTTCAACTT CTTACAGAAC AAATAAGCTC

-2681 -2661 -2641 -2621 -2601
 AAGTCTCTTA TTCCAATTCT TGATCTGTTA CTGCAATAA GGCACCTTGC TTAGATGTAC CGGATGCTTA TGAATTACTG AGTAGGTTAA CTTTAACCGG
 TTCAGAGAAT AAGGTTAAGA ACTAGACAAT GACGTTTATT CCGTGAAACG AATCTACATG GCCTACGAAT ACTTAATGAC TCATCCAATT GAAATTGGCC

-2581 -2561 -2541 -2521 -2501
 GTTTTATCGT CATTAAACCG GAGAAATICA TCTAGTAACC AAATGCTCTG CTGGACCTTT CTTTCAGTGA GCAACTATAG GTGGGTTTTT GGCAGTTGAT
 CAAAATAGCA GTAATTGTC CTCTTTAAGT AGATCATTGG TTTACGAGAC GACCTGGAAG GAAAGTCACT CGTTGATATC CACCCAAAAA CCGTCAACTA

-2481 -2461 -2441 -2421 -2401
 GTACCATAAT TGGTGCAAAC ACACATTTT CTTGAATTTT TGTTTAACTT AAATAAAGTT ACTTCGTTTT CTTGTTTTTT TTAATATGAA TAAAAAAAT
 CATGGTATTA ACCACGTTTG TGTGTAAGAA GAACCTAAAA ACAAAATTGAA TTTATTTCAA TGAAGCAAAA GAACAAAAAA AATTATACCT ATTTTTTTTA

-2381 -2361 -2341 -2321 -2301
 CAACCATAAC TGATAGTAGG TTGTTATCT TTATCAAAAC AAATAAGTTT AATAGGCAGA AAAATAATTG TCTATAGAAT CAATTATGAA AATGCCATTT
 GTTGGTATTG ACTATCATCC AACCAATAGA AATAGTTTGT TTTATTTCAA TTATCCGTCT TTTTATTAAC AGATATCTTA GTTAATACTT TTACGGTAAA

-2281 -2261 -2241 -2221 -2201
 TTTGGGATGG CATTTGTGGA TTTTGCCCTT TTTTAAATAG TTTGTGAATT TTGCCATTTT TCAGGTTACG TGAATGAATA TACGTTTTAT TCATTATGTT
 AAACCCTACC GTAACACCTT AAAACGGGAA AAAAATTATC AAACCTTAA AACGGTAAAA AGTCCAATGC ACTTACTTAT ATGCAAAATA AGTAATACAA

-2181 -2161 -2141 -2121 -2101

TGGGTTTACT CGGTGTGGT TGTCTTAGG GTTAGTATT TTGTGTAAC TACGTATTTT TACCAAAAA AGTCCGAAAT CCATATATTT TTAATCTTA
ACCCAAATGA GCCAACACCA ACAAGAATCC CAAATCATAA AACACATTTG ATGCATAAAA ATGGTTTTTT TCAGGGCTTA GGTATATAAA AATTTAGAAT

5

-2081 -2061 -2041 -2021 -2001
GAAAATGGCT TATCCGTAAG ATTTTAGTAA AAATGGCAAT TTCAAAGAT CTCTATAAAA AATGGCAAAA TCAACAATAA TCCCTTGTCT ATATGGTGGT
CTTTTACCGA ATAGGCATTC TAAATCATTT TTACCGTTA AAGTTTCTA GAGATATTTT TTACCGTTTT AGTTGTATT AGGGAACAGA TATACCACCA

10

-1981 -1961 -1941 -1921 -1901
ATTTCGTCTA AAAGTGAATT ATGGGTAGAT TTTTGTGCTT CATAGATTCT TTGTCGAAAA AAAATTACTT TGTACATTTT AGTGGAGTTA TTTAAATTTT
TAAAGACGAT TTTCACTGAA TACCCATCTA AAAATCGAA GTATCTAAGA AACAGCTTTT TTTTAATGAA ACATGTAAAA TCACCTCAAT AAATTTAAAG

15

-1881 -1861 -1841 -1821 -1801
CCAAATGAAC AAAACCATAT ATTGATGAAA TTGCGAAATG CAATCCAAAA ATAAATATGT TCCACTCTTT TGGTTAGCTT TTAACATAAG ATGCGTTTAA
GGTTAACTTG TTTTGTGATA TAACTACTTT AAGCGTTTAT GTTAGGTTTT TATTATACA AGGTGAGAAA ACCAATCGAA AATTGATTTT TACGCAAAAT

20

-1781 -1761 -1741 -1721 -1701
CTTTATGTAA GTGGTGTATC TTTTGGCAAT GGGGGACAAT GACTATACAA TCTAAGAGAT CATTTTAAAG AATATCATTC ATATTTTCATC CTCTTCTTCA
GAAATACATT CACCAACTAG AAAACCGTTA CCCCTGTATA CTGATATGTT AGATTCTCTA GTAAAAATGC TTATAGTAAG TATAAAGTAG GAGAAGAAGT

-1681 -1661 -1641 -1621 -1601
AAATTCAGTT TCACTAATTA ACCACGTTTC AATTGTAGTG TATCGCGAGC TGTAATATTT ATCTAATTTA TGTACATAAA TCATAACTGT AATCTTTTAT
TTAAAGTCAA AGTGATTAAT TGGTGCAAA GTTAATCATC ATAGCGCTCG ACATTTATAA TAGATTAAAT ACAATGTATT AGTATTGACA TTAGAAATAA

-1581 -1561 -1541 -1521 -1501
AGACAAAAAC ATATATACCT CACTGCAAAC ACCTTCAAAC ATGGATAACT TGATTTAGGC ATACAAATAT TATTCTCAT TTTTGTGATA TGACCTATAT
TCTGTTTTTG TATATATGGA GTGACGTTTG TGGAAATTTG TACCTATGTA ACTAAATCCG TATGTTTATA ATAAAGAGTA AATAACTAT ACTGGATATA

-1481 -1461 -1441 -1421 -1401
TATGTGGCTA TTTTATCAGT TTTAGTGT TTATGATAA TTGAACCACT TAAATGTTT TCTCATTTTT CAATTTATTT TAAACTGAAT TAAAAAGTAA
ATACACCGAT AAAATAGTCA AAATCACAAA AAATACTATT AACTTGGTGA ATTTACAAAT AGAGTAAAA GTTAAATAA ATTTGACTTA ATTTTTCATT

-1381 -1361 -1341 -1321 -1301
GAAAGTATGA TCCAAATAGG CATCGACACA TGGAAACCCA TTTTAAGSTA GAAGATGCTT TTCTCGGCT TCTGAAACA ACTAGAAAAT GATATGATAC
CTTTCATACT AGGTTATTCC GTAGCTGTGT ACCTTTGGGT AAAATTCAT CTCTACGAA AAGACGCCGA AGACTTTTGT TGATCTTTTA CTATACTATG

-1281 -1261 -1241 -1221 -1201
GTTGCTTTCA TTTATGTAA GTATTATTTA GTTTTAAATC ACCGCTTCA TATCCAGCTG CAAGACTACT ACAACTTGCA ATTATGAGAC TCTCGTTAGA
CAACGAAAGT AAATAACATT CATAATAAAT CAAAATTAAG TGCGCGAAGT ATAGGTGCGAC GTTCTGTAGA TGTGAAAGT TAATACTCTG AGAGCAATCT

-1181 -1161 -1141 -1121 -1101
AAATTACCAG GTATAATTTA AAAACAAAA GAACTAGAAT ATATTGGCAA TTATTGTAAG TAAGAAAATA TGAGATTCTT GACCGAGTTG TTAACATATC
TTTAATGCTC CATATTAAAT TTTTGTTTTT CTGTACTTTA TATAACGGTT AATAAACTTC ATTCTTTTAT ACTCTAAGAA CTGCTCAAC AATTTGATAG

-1081 -1061 -1041 -1021 -1001
AAACCCAAAA GTTTTGGTTA AAAAATAAGC TAGTACTATG TACATATGTT TTATGTTGAA AATATATTAA ACTGTATGTA AGAGGGAGTG TACTTTTCAAT
TTTGGGTTTT CAAAACCAAT TTTTATTTCG ATCATGATAC ATGTATACAA AATAACAATT TTATATAATT TGACATACAT TCTCCCTCAC ATGAAAGTAA

-981 -961 -941 -921 -901
TTAGATATAC ATTTCCAGCT AGTACGAGGT CTCTATATAT AAATTTCTT AATATCGCTA AACAAATTTT ACTTTCAAGT TTGTAATGTG ATAAATGAAA
AATCTATATG TAAAGGTGCA TCATGCTCCA GAGATATATA TTGAAAGAA TTATAGCGAT TTGTTTAAAA TGAAGATTCA AACATTACAT TATTCACATT

-881 -861 -841 -821 -801
GACCGTATAT ACATACACAT GTTAATCAAC TGATAACCTT TGTGCTCGT GTGCTAGTT ACTAGTCAAC CATCAACGT GCATGATGCT GTTTTCTTA
CTGGCATATA TGTATGTGTA CAATTAGTTG ACTATTGGAA ACACGGAGCA CACAGATCAA TGATCAGTTG GTAGTTTGCA CGTACTACGA CAAAAAGAA

-781 -761 -741 -721 -701
GAGTACTATT GTTGTGTTAT ATATAACTAA ACATAAACAA TTGTCTATTA TGATATAAAC ATAGAATTTT CAAGCAATGA TATGTTTGA TGTTTGTAT
CTCATGATAA CAACACAATA TATAITGATT TGTATTGTIT AAACGATAAT ACTATATTTG TATCTTAAAA GTTCGTTACT ATACAAATCT ACAAACATA

-681 -661 -641 -621 -601
AAATATTCCA TAAATAGTAG ACACCCATAT ATACACAAAC ATGAATTTCTA CCTGAGGAGA AACACATAGA TGTTCAAATT AAATAATAAC CCTATAATGA
TTTAAAGGT ATTTATCATC TGTGGGTATA TATGTGTTG TACTTAAGAT GGACTCCTCT TTGTGTATCT ACAAGTTTAA TTTATTATTG GGATATTACT

-581 -561 -541 -521 -501
AAACTCTAAA GTAAGTAATA CGAAATAAAA ATTTATCTCT TAAATAACAT ATAAACATAT ATATACAAGT TTAATTTGTA ATTGTATCAC AAGAGCCAA
TTTGAGATTT CATTCATTAT GCTTTATTTT TAAATAGGAA ATTTATTGTA TATTGTGATA TATATGTTCA AATTAACCAT TAACATAGTG TTCTCGTTA

-481 -461 -441 -421 -401

TATTTGGTGA CTGTATCACA CGTGCTTAAA GAGAGCGTGG GAATGAAAGT AAAGAAGAAT AAAGAAGCAG AGAGATGGGC TAGAAATGAG AAAACACACC
ATAAACCACT GACATAGTGT GCACGAATTT CTCTCGCACC CTTACTTCA TTTCTTCTTA TTTCTTCGTC TCTCTACCGG ATCTTTACTC TTTTGTGTGG

-381 -361 -341 -321 -301
AAACCCTAAC CTCACCTCA CACATTTCTT ATCTTTTGCT CTCAATAGAT TCCATTGATT CAAAACAAA TTTTCATTAA GATTTACAA CCTCCACACA
TTTGGGATTG GAGTGGGAGT GTGTAAAGAA TAGAAAAAGA GAGTTATCTA AGGTAACATA GTTTTGTITT AAAAGTAATT CTAAAGTGTT GGAGGTGTGT

-281 -261 -241 -221 -201
CTTCCAAACA CAATTAAAGA GAGGAAAAAG AATCAATAAC CCTATAATA AAAATCAGA CAAACAGAAG TTTCTCTTC TTCTTCTTA AGCTAGTACC
GAAGGTTTGT GTTAATTTCT CTCCTTTTTC TTAGTTATTG GGATATTTAT TTTTGTAGTCT GTTTGTCTTC AAAGGAGAAG AAGAAGGAAT TCGATCATGG

-181 -161 -141 -121 -101
TTTTGTCTT GAAATTAGG TTAATTTCTT TTTTCCAAAT ACCATCAATT CTCCAGACCA TAAAACTCA AAAAGATCAG ATCTTTCTCT TGAAGAGAG
AAAAACAAGAA CTTTAATCCC AATTAAAGAA AAAAGGTTTA TGGTAGTTAA GAGGTCTGGT ATTTTGTAGT TTTTCTAGTC TAGAAAGGAG ACTTTTCTCT

-81 -61 -41 -21 -1
ATACCAACT TATGTTTTTG TGTGTCTGTA TATAGATAAA CATTACATAC CCATATTTGT GTATAGACAT AAAAAGTGGA AATTAGGTA AAAAAAGAA
TATGGGTTGA ATACAAAAAC ACACAGACAT ATATCTATTT GTAATGTATG GGTATAAACA CATATCTGTA TTTTTCACCT TTAATTCAT TGTTTTCTT

20 40 60 80 100
ATGGAAGAG GAAGAGTAGA GCTGAAGAG ATAGAGAACA AAATCAACAG ACAAGTAACG TTTGCAAGC GTAGGAACGG TTTGTTGAAG AAAGCTTATG
TACCTTCTCT CTCTCATCT CGACTTCTCC TATCTCTGT TTAGTTGTC TGTTCATTGC AAACGTTTCG CATCTTGCC AAACAACATC TTTCAATAC

120 140 160 180 200
AATTGTCTGT TCTCTGTAT GCTGAAGTTG CTCTCATCAT CTTCTCCAAC CGTGGAAGC TCTATGAGTT TTGCAGCTCC TCAAGTAAA CAACTCTCTC
TTAACAGACA AGAGACACTA CGACTTCAAC GAGAGTAGTA GAAGAGGTTG GCACCTTTCG AGATACTCAA AACGTCGAGG AGTTTCAATT GTTGAAGAG

220 240 260 280 300
ACTCTTTATC AGTTTCTTGA TTGAGTTTTT GCTAGATCTG AGCTTAGATC TTTGTCTCAA GGAATTGTTA TATATAGATC ACACGATCTT GATTCTACG
TGAGAAATAG TCAAGAAGCT AACTCAAAAA CGATCTAGAC TCGAATCTAG AAACAGAGTT CCTGAACAAT ATATATCTAG TGTGCTAGAA CTAAGATGC

320 340 360 380 400
AAGTTGAGTT AATTAGATT CTTGATTCA TTTTCTAGGG TTTTCTTCCA ATTCTTGAAA TTTAAGATCT GGTTTTTTTG TTGTCAATGA TTTGAACTG
TTCAACTCAA TTAATCTAAA GAACTAAAGT AAAAGATCCC AAAAAAGGT TAAGAACTTT AAATTCTAGA CCAAAAAAC AACGTTACT AAATCTTGAC

420 440 460 480 500
TGAATTTTGT AATCGAATAG ATTCCAAATC CTGATATGCA ATCTGAAAAG TTTTATATAA TTAATATATG TCTGTGTGAT TGGAACTTA AAAGTTGTTT
ACTTAAACA TTAGCTTATC TAAGGTTTAG GACTATACGT TAGACTTTTC AAAATATATT AATTATATAC AGACACACTA ACCTTTGAAT TTTCAACAAG

520 540 560 580 600
ACAGATTCT ATGAAAATTA CAAGTATCCA ACGTAGAATG ATAATATATG GTTACATGCA TTAACCATTT GITAGTTTAT CATACTTAT GGTGGTTAAA
TGTCTAAGA TACTTTTAT GTTCATAGGT TGCATCTTAC TATTATATAC CAATGTACGT AATTGGTAAA CAATCAAGTA GTATGAAATA CCACCAATTT

620 640 660 680 700
ACTTCAACG CGTGATATC TGTGAAGGCT TTGATTGTTT GTTTTTCTT AAAAAAATG TTTAATAGAT TTTAATTAT ATGTTAAAT AGTTTGTCTT
TGAAGTTTGC GCACATATAG ACACCTCCGA AACTAACAAA CAAAAAGAA TTTTGTGTTT AAATTATCTA AAAATTAATA TACAATTTTA TCAAAACGAA

720 740 760 780 800
ACATGCATTC AAGAAAATAT AGCGATTAT TCCTTTTTTC AAATCACAAT TTGTGAATCA AACGAAAACG TAAGATATTG CTGCAAAATG ATAGGATTGA
TGTACGTAAG TTCTTTTATA TCGCTAATTA AGGAAAAAG TTAGTGTTA AACACTTAGT TTGCTTTTGC ATCTATAAC GAACGTTTAC TATCTAACT

820 840 860 880 900
ACTATGATA TTTGTAATA TAAATACGAA ACTTTACGTT TGAAAGTTGA AACAATCAAA TCCAAATCAA CTCGTATATA ATCAGATAAA TAATGGAAAC
TGATAACTAT AAACATTTAT ATTTATGCTT TGAATGCAA ACTTTCAACT TTGTTAGTTT AGGTTTAGIT GAGCATATAT TAGTCTATTT ATTACCTTTG

920 940 960 980 1000
AATCTTCAAT TTTGATGGAA GAATACTTTA AAACITGAAG AGCTTTTTTT TTAGGTGAT TTATAGGTTT AGATCTCCAA AGTCAAGTAT GATCTTTTTA
TTAGAAAGTTA AAACATCTT CTTATGAAAT TTTGAACITC TCGAAAAAA AATACCACTA AATATCCAAA TCTAGAGGTT TCAGTTTATA CTAGAAAAAT

1020 1040 1060 1080 1100
ATAAACTCTT ATCTCTCTT TTTGAGTTAT TTTGAGTATG CTCAAGACAC TTGATCGGTA CCAGAAATGC AGCTATGGAT CCATTGAAGT CAACAACAAA
TATTTGAGAA TAAGAGAGAA AAACCAATA AAGTCGTAT GAGTCTGTG AACTAGCCAT GGTCTTTACG TCGATACCTA GGTAACTTCA GTTTGTGTTT

1120 1140 1160 1180 1200
CCTGCCAAAG AACTTGAGGT GTTCTTAATT CAAATACTAT TTTGATTCC TATCATATCA TTTCAAGAAA GATCTTTTTT AAAAGTTTGT TTTCTGTAAG
GGACGGTTTC TTGAACCTCA CAAGAATTAA GTTTATGATA AAATCTAAGG ATAGTATAGT AAAGTTCTTT CTAGAAAAAA TTTTCAACA AAAGCACTTT

1220 1240 1260 1280 1300

TATTTTCAGAA CAGCTACAGA GAATATCTGA AGCTTAAGGG TAGATATGAG AACCTTCAAC GTCAACAGAG GTACATATCT GTCTACCTCC GTATATTTAC
ATAAAGTCTT GTGATGTCT CTATAGACT TCGAATTCCT ATCTATACTC TTGGAAGTTG CAGTTGTCTC CATGTATAGA CAGATGGAGG CATATAAATTG

1320 1340 1360 1380 1400
TCAATTCTGT ATCCATGTAG ATTATATTT GTAGGTGTGT GTGGCTTTTG TTGGTGCGA AATCTTCTTG GGGAGGATT AGGACCTTTG AATTCAAAGG
AGTTAAGACA TAGGTACATC TAAGTATAAA CATCCACACA CACCGAAAAC AACCAAGTCT TTAGAAGAAC CCTTCCTAAA TCCTGGAAC TTAAGTTTCC

1420 1440 1460 1480 1500
AGTTAGAGCA GCTTGAGCGT CAACTGGAGG GCTCTCTCAA GCAAGTTCGG TCCATCAAGG TATCTTTATA CATGGAATCA ATGATTCAAA TGAGATTAAT
TCAATCTCGT CGAAGTCGCA GTTGACCTGC CGAGAGAGTT CGTTCAAGCC AGGTAGTTCC ATAGAAATAT GTACCTTAGT TACTAAGTTT ACTCTAATTA

1520 1540 1560 1580 1600
TTGTGTGTGT TAATTATAAC TACTATGGTG GTATGATGAT TGTTTGCAGA CACAGTACAT GCTTGACCAG CTCTCGGATC TTCAAAATAA AGAGCAAATG
AACACAACAA ATTAATATTG ATGATACCAC CATACTACTA ACAAACGTCT GTGTCAATGA CGAACTGGTC GAGAGCCTAG AAGTTTATT TCTCTTTAC

1620 1640 1660 1680 1700
TTGCTTGAAA CCAATAGAGC TTGGCAATG AAGGTATAAT TACAGAAATA ATGCATTGG TGCCCTGCGA TCAATCTCTT TCACAGAGTT TAAGTTTCTA
AACGAACCTT GGTATTCTCG AAACCGTTAC TTCCATATTA ATGCTTTATT TACGTAAACC ACGGAACGCT AGTTAGAGAA AGTGTCTCAA ATTCAAAGAT

1720 1740 1760 1780 1800
AACATTTTGG GAAACATCTC TAGTTTTCTT GTTCTGATT ATAGTCTTTT GGTGAAATGT AAATGTTTAG CTGGATGATA TGATTGGTGT GAGAAGTCAT
TTGTAAAAAC CTTGTAGAG ATCAAAAGAA CAAAGACTAA TATCAGAAAA CCACTTTACA TTACAAATC GACCTACTAT ACTAACACA CTCTTCAGTA

1820 1840 1860 1880 1900
CATATGGGAG GAGGAGGAGG ATGGGAAGGT GGTGAACAGA ATGTTACCTA CGCGCATCAT CAAGCTCAGT CTCAGGACT ATACCAGCCT CTTGAATGCA
GTATACCCCT CTCTCTCTCC TACCCTTCCA CCACTTGTCT TACAATGGAT CGCGTAGTA GTTCGAGTCA GAGTCCCTGA TATGTGCGGA GAACCTACGT

1920 1940 1960 1980 2000
ATCCAACCTC GCAATGGGG TAAATCCTTT GCCTTAAACA ATCATCTGCA AATCAGCTTG TGTACTTCAC TACTAAGATT GTACTTATAT AAGTCTCTT
TAGGTTGAGA CGTTTACCCC ATTTAGGAAA CGGAATTTGT TAGTAGAGCT TTAGTCGAAC ACATGAAAGT ATGATTCTAA CATGAATATA TTCCAAGAAA

2020 2040 2060 2080 2100
AGTTACTTGG TGTAAGAGG ATCATCAATG TGTGTGAACC TTTTAAGTTG CTGTTTTGGT GATGATGATG ATGATGACAG GTATGATAAT CCGGTATGCT
TCAATGAACC ACATTTCTCC TAGTAGTTAC ACACACTTGG AAAATTCAAC GACAAAACCA CTACTACTAC TACTACTGTC CATACTATTA GGCCATACGA

2120 2140 2160
CAGAGCAAA ATCTGCGACA ACCCAAGCTC AGGCGCAGCA GGGAAACGGT TACATCCCGG GGTGGATGCT C
GTCTCGTTA TTGACGCTGT TGGGTTCGAG TCCGCGTCTG CCTTTTGCCA ATGTAGGGCC CCACCTACGA G

SEQ ID NO:42

Arabidopsis SEP2 genomic sequence

-2981 -2961 -2941 -2921 -2901
ACGCTCTAAC CAACTGAGCT AATGGGCCAT TTGCGAATGG TAGTGTCTAT TTTACTTATT CGAATCTAAA TCGTCATAGG TAATTAAGAA GACATGCAAA
TGCGAGATTG GTTGACTCGA TTACCCGGTA AACGCTTACC ATCAGAGATA AAATGAATAA GCTTAGATT AGCAGTATCC ATTAATTCTT CTGTACGTTT

-2881 -2861 -2841 -2821 -2801
GCTTAATCAA TGATGGATC TTGATTCTA CTCTAGGTG CCACCATTGA CGCATTCTA AAATCATAAC CGGTCGTTTA CAAACATAT TGCTTGAATG
CGAATTAGTT ACTACCTAAG AAACCTAAGT GAAGATCCAC GGTGGTAAT GCGTAAGTAT TTTAGTATTG GCCAGCAAT GTTTTGTATA ACGAATTAC

-2781 -2761 -2741 -2721 -2701
ATTCTAAACA AATAATAGTT TTTTGTGAA ATTTTCAAAA CATATGTTAG GTAAGGTGAG GTTTTGCCAA TAAGCCTTAC TATATACAGT GGCAACATGT
TAAGATTGTT TTATTATCAA AAAACAACCT TAAAAGTTT GTATACAATC CATTCCAGTC CAAAACGGTT ATTCGGAATG ATATATGTCA CCGTTGTACA

-2681 -2661 -2641 -2621 -2601
TTCTTCTACT TTGGAGGATT TTGGGTGAAT ATGAAACCCA TGTGAGCATG ATACATGTTT TTCTTCTTCT ATTGAAATTT CCCCAGTGG TCATTGTGCTC
AAGAAGATGA AACCTCTTAA AACCCACTTA TACTTTGGGT AACTCTGTAC TATGTACACA AAGAAGAAGA TAACTTTAAA GGGGGTTACC AGTAAACGAG

-2581 -2561 -2541 -2521 -2501
TTTGCGTTTG TGTGCGCTT TCCGGTATCA AATCATATAT ATATATAACC TAAATGAGAC TAGACAATTT GAATCATTTT AAAAGGTATA AAGAAGATG
AAACGCAAGC ACAACGCGAA AGGCCATAGT TTAGTATATA TATATATTGG ATTTACTCTG ATCTGTTTAA CTTAGTAACA TTTTCCATAT TTCTTCTCTA

-2481 -2461 -2441 -2421 -2401
TATAGTCCAC AATTAAACAA GTAATAAGAC GGTAAATAT CAAACAAAT GAAAGGGTAA AAAAAAACA AGAGGGACAA GTCACTGTTA GAAAGGTGAC
ATATCAGGTG TTAATTGTTT CATTATCTG CCATTTTATA GTTTGTTTAA CTCTTCCATT TTTTTTTGT TCTCCCTGTT CAGTGACAAT CTCTCCACTG

-2381 -2361 -2341 -2321 -2301
TCCTCCCTTT GGGCCAGCCC CCTACCAAA AAGTCAAAGC TTACTTACTA TTCAGTCATA TATGACACG TGTACTTCCA ACCACATCAC CCATCTCTTT

AGGAGGGAAA CCCGGTCGGG GGATGGTGT TTCAGTTTCG AATGAATGAT AAGTCAGTAT ATAGCTGTGC ACATGAAGCT TGGTGTAGTG GGTAGGATAA

-2281 -2261 -2241 -2221 -2201
ACGTAATTTC CACTGTCTAG ACTTTTTTTT TTTTTTTTTT TTTTACTTTT TAACGTTTTT TAGCTGTCTC TCTAAATTAC TACATACGGA CTTGTCTACGT
TGCATTAAAG GTGACAGATC TGAAAAAATA AAAAAAATAA AAAATGAAAA ATGTCACAGAG AGATTTAATG ATGTATGCCT GAACGATGCA

-2181 -2161 -2141 -2121 -2101
CACCTGAGAA GAAAGATCTT TGCTCGTAGA TTCTTTGTCT GAAGGAAAAAT TATTGTGATT TAGTTATTTA CAATTGCATA ATTGTGTGTA GTAAATCCGC
GTGGACTCTT CTTTCTAGAA ACGAGCATCT AAGAAACAGA CTTCCTTTTA ATAAACATAA ATCAATAAAT GTTAACGTAT TAACACACAT CATTTAGGCG

-2081 -2061 -2041 -2021 -2001
CAGAATGATA TTAGAGTGAT ACTGAGACGA CGAATGGTGT AACTTGTAAC ATATATACTA ATAAACACGA TTGATTAAAA ATTTACTATA CAGTATATCC
GTCTTACTAT AATCTCACTA TGACTCTGCT GCTTACCACA TTGAACATTG TATATATGAT TATTGTGCT AACTAATTTT TAAATGATAT GTCATATAGG

-1981 -1961 -1941 -1921 -1901
AAAAATTAT GATTGAGAGT GTACATATAC AATAAGTAAT TAAACCTCAA AACCACACAG TTTTTTTTTT TTTTGGTCAA CAATAATTAG AAATGAGAAT
TTTTGTAATA CTAATCTCA CATGTATATG TTATTCATTA ATTTGGAGTT TGGTTTGTG AAAAAAATAA AAAACCAAGT GTTATTAAAT TTTTACTCTA

-1881 -1861 -1841 -1821 -1801
AAACTATTTA ACTTATAAAT TGTAGACCCA AAAACTCATA TTTTACCCTT CTTGGTCTCA CCTAAAAAGA CTTTAAATCC CAAAACCTCT GCAAAACAATG
TTTGATAAAT TGAATATTTA AGATCTGGGT TTTTGAGTAT AAAATGGGAA GAACCAGAGT GGAATTTTCT GAAATTAAGG GTTTTGAGAA CGTTTGTATC

-1781 -1761 -1741 -1721 -1701
GCCAAACATA GAAGATTGGA AAACAAATTT AAATCTACTT TCACCTTTTAT AAAGAATAAT CAACGAACCA ATTAAGTTAA ACCTACATAT ATTCGTATGT
CGGTTGTAT CTTCTAACCT TTTGTTTAAA TTTAGATGAA AGTGAAATAA TTTCTTATTA GTTGTCTGGT TAATTCATTT TGGATGTATA TAAGCATACA

-1681 -1661 -1641 -1621 -1601
GATCACATAT GTGTTATATT CCTCACGTC TCTTCCATT AGCTAATAAC CTTAATTACT TCAAGAAATC ATATATCAAC CGAAACTAG TAAATAAAT
CTAGTGATA CACAATATAA GGAGTGCAAG AGAAGGTAAA TCGATTATTG GAATTAATGA AGTCTTTAG TATATAGTTG GCTTTTGATC ATTTTATTTA

-1581 -1561 -1541 -1521 -1501
ATACATACTG AAAGCGCGCA AAATTTTATG CAATATTTTA AAATACCTTA CATCATAGTC TTAACATAAT AATCTTTCTG ATCAAAATTT ATTTTCATAA
TATGTATGAC TTTGCGCGCT TTTAAAAATC GTTATAAAT TTTATGGGAT GTAGTATCAG AATTGATTAA TTGAAAAGAC TAGTTTAAA TAAAGTATT

-1481 -1461 -1441 -1421 -1401
TATTCATAAA TACTTATGGA TTACCTAAAC CAGGATACCT ATCCCTATAA ATCTGTCAAT CATCATGGAT TCATGGAGAC ATGGTCAGAT ATCCCACGTC
ATAAGTATT ATGAATACCT AATGATTG GTCTATGAA TAGGATATT TAGACAGTTA GTAGTACCTA AGTACCTCTG TACCAGCTA TAGGGTGCAG

-1381 -1361 -1341 -1321 -1301
CAGATACAAT GTAACATATT GATATACTGC GGCTGATTAT TATTTTITAC ATTAGAACGA GTTTAGATCC AAAACAAAAT TGGTATTCTC AAACAAAAT
GTCTATGTTA CATGTATATA CTATATGACG CCGACTAATA ATAAAAATG TAATCTTGCT CAATCTAGG TTTTGTTTTA ACCATAAGAG TTTTGTTTTA

-1281 -1261 -1241 -1221 -1201
TAAAAATTGA ATACGAAGT AATAGACAA AACTTCAATG TTGTCGAATA GATAGGAAGC AATAGAAAAG CGACACGTAC ATGTCCATTT TAAGGTAGGA
ATTTTAACT TATGCTTTCA TTATCTTGT TTGAAGTTAC AACAGCTTAT CTATCTCTCG TTATCTTTTC GCTGTGCATG TACAGGTAAA ATCCATCCT

-1181 -1161 -1141 -1121 -1101
GAGGCTTTTC TGCGGCTTGT GAAGTAAGAA AAAGAAAATG ATGATAGCTG CTTTGTGTTT ATTCATTGCA GAAGAAACCA ATGTPTCCCC AATCTCACGC
CTCCGAAAAG ACGCCGAACA CTTCATTCTT TTTCTTTTAC TACTATCGAC GAAAGCAAAG TAAGTAACGT CTTCTTTGGT TACAAAGGGG TTAGAGTGCG

-1081 -1061 -1041 -1021 -1001
GCCTCTCTCT ATCTACCACC ACTTGACAA ATCCCTTTT CAGTATTAGT TTTTITTTCC GGACATTGTA CATTCAAAAG CATTCCAAGT GTCTAATAAA
CGGAGGAGGA TAGATGGTGG TGAACCTGTT TAGGGGAAAA GTCATAATCA AAAAAAAGG CCTGTAACAT GTAAGTTTTC GTAAGGTTC AAGATTATT

-981 -961 -941 -921 -901
CATAACTAAC CACTCCAAGA TGCAAAATCT AGCTACGAAC AAATTTTAAA CTATAGAGAT GAACCTTAAA TTCGGGCATT AATTAGTGGA ACTTGAGCTA
GTATTGATTG GTGAGGTCT ACGTTTGTAG TCGATGCTTG TTTAAAATTT GATATCTCTA CTTGAAATTT AAGCCCGTAA TTAATCACCT TGAATCGAT

-881 -861 -841 -821 -801
TTGATGAGTT TTCTGACTTT TTGAAGCTTA ATTGAGTTTT ATATACACTA TATATAGGCT TGTAATAATA TGGATCAAA AAGAAATATA TAAACTACAA
AACTACTCAA AAGACTGAAA AACTTCGAAT TAACTCAAAA TATATGTGAT ATATATCCGA ACATTATTAT ACCTAGTTTG TTCTTTATAT ATTTGATGTT

-781 -761 -741 -721 -701
ATTGGGAATT AGGTTTAAA ACGTTATCGT TCTATTTTAA TTCAGGCACC TTTAGAATAT CAAGATCCAT GCATGTTTCA ATATTCTGT TGACAAATAA
TAACCTTAA TCCAAAATTT TGCAATAGCA AGATAAAAT AAGTCCGTGG AAATCTTATA GTTCTAGGTA CGTACAAAGT TATAAGACA ACTGTTTATT

-681 -661 -641 -621 -601
ATAAAGATGT CTCAATATG AAGTTTGGGC AACGTACGTG TAGACCTAAA AGAGTCGAAA CATTGGTATC TAAGTCATAT ATCTAGATGT ATATGGACAT

5
10
15
20
25
30
35
40
45
50
55
60
65

TATTTCTACA GAGTTTATAC TTCAAACCCG TTGCATGCAC ATCTGGATT TCTCAGCTTT GTAACCATAG ATTCACTATA TAGATCTACA TATACCTGTA

-581 -561 -541 -521 -501
GGATTATATA ACTAGACAAC GTTTGTTTAA AAACTTAAAT TCATTTTCT TAATTAGTAG CAACCTAGCAA CTAACCTACTC ATGGCAAATA ATGGTGTCTG
CCTAATATAT TGATCTGTG CAAACAAAAT TTTTGAATTA AGTAAAAAGA ATTAATCATC GTTGATCGTT GATTGATGAG TACCCTTTAT TACCACAGAC

-481 -461 -441 -421 -401
CGTGGCAGCG ACTTGGGAGA GAAGGTGTGA GAATGTTTIT TACTTTCTGT GTAAAAGATG GAAGAGAGAG AAAGAGTAAA GAAGTAGAGA GAGAGATATT
GCACCGTGCG TGAACCCCTCT CTTCCACACT CTTACAAAAA ATGAAAGACA CATTTCTAC CTTCTCTCTC TTTCTCATTT CTTCTCTCTC CTCTCTATAA

-381 -361 -341 -321 -301
GTATCACCAA ACCCTAATGA TCTCTACCCC TCACAAATTT TCTTATCTTT ATAGCTTTTA TAGATTACAA AAACTTTTC TTCAGATTCA CAATCTCATC
CATAGTGGTT TGGGATTACT AGAGAGTGGG AGTGTTTAAA AGAATAGAAA TATCGAAAAT ATCTAAGTGT TTTTGAAAAG AAGTCTAAGT GTTAGAGTAG

-281 -261 -241 -221 -201
ACAACCCCTC AAAAAGAGAA AAGATCTAAA GAATAAACAA GAGCCCTAAT ATCAATCAC AACCAAAAAA ACCAAAGAAA GCTAATTAAA GTTTTCTCTC
TGTGGGAAG TTTTCTCTTT TTTTAGATTT CTTATTTGTT CTCGGGATTA TAGTTTAGTG TTGGTTTTTT TGGTTTTCTT CGATTAATTT CAAAAGAGAG

-181 -161 -141 -121 -101
TAGCTATTC TCTCTTTTC TTGTTCTTGA AAATCAGGGT TTACTTCACC AAAAGATAAG ATCTTTCCCC AGAAAAAGCA ATACCCCAAGT CATGTTTCTG
ATCGATAAG AGAAGAAAAA AACAAGAACT TTTGATCCCA AATGAAGTGG TTTTCTATTC TAGAAAGGGG TCTTTTTCGT TAGGGTTCA GTACAAAGAC

-81 -61 -41 -21 -1
TGTGTCTGTA TATAGATAAA ACATTACATA CCCTAATAAG GTTACACAAA TAGCTATAAA AGAGGGAAAA TAAGATAGGG ATTTTGTGG GTGAGGAAAG
ACACAGACAT ATATCTATT TGAATGTAT GGGATTATTC CAATGTGTTT ATCGATATTT TCTCCCTTTT ATTCTATCCC TAAAAACCC CACTCTTTTC

20 40 60 80 100
ATGGGAAGAG GAAGAGTAGA GCTCAAGAGG ATAGAGAACA AAATCAACAG ACAAGTGACG TTTGCTAAAC GTAGAAATGG TTTGCTGAAA AAAGCTTATG
TACCCTTCTC CTTCTCATCT CGAGTTCTCC TATCTCTGT TTTAGTTGTC TGTTCACTGC AAACGATTG CATCTTTACC AAACGACTTT TTTCAATAC

120 140 160 180 200
AGCTTTCTGT TCTCTCGAT GCTGAAGTCT CTCTCATCGT CTTCTCCAAC CGTGGCAAGC TCTACGAGTT CTGCAGCACC TCCAAGTACT TCTCTTTCTT
TCGAAAGACA AGAGAGCCTA CGACTTCAGA GAGAGTAGCA GAAGAGGTTG GCACCGTTCC AGATGCTCAA GACGTCGTGG AGGTTCTAGA AGAGAAAGAA

220 240 260 280 300
TATACACTTA TTAGATCTGT GTGTAGATCT TTCATTTTC TAGTCTTGTG ATGAGTTTTA TCTTTCTTGA TTGCTTTTTA ACAAATACT TGATATATTT
ATATGTGAAT AATCTAGACA CACATCTAGA AAGTAAAAAG ATCAGAACAC TACTCAAAAT AGAAAGAACT AACGAAAAAT TGTTTTATGA ACTATATAAA

320 340 360 380 400
TCAGTTTCTT AATCTAGCT CTAATTAGGT TTTGATTATA GAAGAATAAT TCAGTACTTT CAAGTGATTG AATTTCGAGA TCTGATCTTA ATTTAATCAT
AGTCAAGAA TTAGACTAGA GATTAAATCCA AAATAATAT CTTCTTATTA AGTCATGAAA GTTCACTAAC TTAAGCTCT AGACTAGAAT TAAATTAGTA

420 440 460 480 500
CATGTCAAAT TCTTAGGAT TTAATTGCAA TCTATTTTTA GATTATCGG AGCTAGGAAA GTATCATAAT GATATACTAT TATTATCATG TAATTTTCATT
GTACAGTTTA AGAATCCCTA AATTAACGTT AGATAAAAA CTAAATAGCC TCGATCCTTT CATAGTATTA CTATATGATA ATAATAGTAC ATAAAGTAA

520 540 560 580 600
GTCTCTACAC GGATATATAT GTGATTAGAA CTTGGTAAAG TAACTAAAG ATTCACAGTC TTCAATGAAA TTTAAAAGAT CCAACGTAGA ATAATTAGTG
CAGAGATGTG CCTATATATA CACTAATCTT GAACCAATTC ATTTGATTTC TAAGTGTGAG AAGTTACTTT AAATTTTCTA GGTGTCATCT TATTAATCAC

620 640 660 680 700
GTTCCATGCA TTAACAGTCT TAATTAAAGC TCATGCAGAC ATTTAAGCAC CACATGAATT TAATATCTTT TTAATTAAAG GATCTTCTTT TTATAAATTT
CAAGGTACGT AATTGTCAG ATTAATTTCG AGTACGTCGT TAAATTCGTG GTGTACTTAA ATTATAGAAA AATTAAATTC CTAGAAGAAA AATATTTAAA

720 740 760 780 800
TCTTTTGTGA GTTTTAAAA TTTTAGTTTG TTCAATTAAT TTATAGATTC TTCTCTCTCT TTTTGTGTT TTTTGATCTT TCAGCATGCT CAAGCACTG
AGAAAAAAT CAAAAATTTT AAAATCAAAC AAGTAATTTA AATATCTAAG AAGAAGAGGA CTAACACAAA AAACTAGAAA AGTCGTACGA GTTCTGTGAC

820 840 860 880 900
GAAAGGTATC AGAAGGTAG CTATGGCTCC ATTGAAGTCA ACAACAAACC TGCTAAAGAG CTTGAGGTTT AATCTCCAAC ATCTCTTCGA TCTTAATAT
CTTCCATAG TCTTCACATC GATACCGAGG TAATCTCAGT TGTGTTTGG ACGATTTCTC GAATCCAAA TTAGAGGTTG TAGAGAAGCT AGAATTAATA

920 940 960 980 1000
TTATCTTTT TTAATTTTAT CTAAAGAAAA TGTGTTGATT TGAGACAAAA GCCCTTCAAA GTTCTTACA TAGATATTCA ATTGCTATT ATCTTCGCAA
AATAGGAAAA AATTAAATA GATTCTTTT ACAACTAAA ACTCTGTTTT CGGGAAGTTT CAAAGAATGT ATCTATAAGT TAACAGATAA TAGAAGCGTT

1020 1040 1060 1080 1100
TTTTCAAGAC AGCTACAGAG AGTACTTGAA GCTGAAAGGT AGATATGAAA ATCTGCAACG TCAGCAGAGG TATATACATT AATGTGGATG ATGATCATTT

AAAAGCTTTG TCGATGTC TCATGAACCT CGACTTTCCA TCTATACCTT TAGACGTTGC AGTCGTCTCC ATATATGTAA TTACACCTAC TACTAGTAAA

1120 1140 1160 1180 1200
ATAAACAGCA TATATATATA TATATATATA TATATATATA GTTTGTATTG ATCATGAAAG TGTGTTGCTG CAGAAATCTT CTGGAGAGG ATCTGGGACC
TATTTGTCGT ATATATATAT ATATATATAT ATATATATAT CAAACATAAC TAGTACTTTC ACACAACGAC GTCTTTAGAA GAACCTCTCC TAGAACCTGG

1220 1240 1260 1280 1300
TCTGAATTCA AAGGAGCTAG AGCAGCTTGA GCGTCAACTA GACGGCTCTC TGAAGCAAGT TCGCTGCATC AAGGTGATTT ACTTCTGTAC ATACACTGAA
AGACTTAAAGT TTCCTCGATC TCGTCGAACT CGCAGTTGAT CTGCCGAGAG ACTTCGTTCA AGCGACGTAG TTCCACTAAA TGAAGACATG TATGTGACTT

1320 1340 1360 1380 1400
AGATTACAC AAATCTTTCT CTATATATAG ACTGAGACAC ATGCATGAAA TGTTTTGTAT GCGTGAGGTT ATCTGAAAT GCCTCTTCTT TTTTGCAGAC
TCTAAGTGTG TTTAGAAAGA GATATATATC TGACTCTGTG TACGTACTTT ACAAAACTA CGCACTCCAA TAGACTTTTA CGGAGAAGAA AAAACGTCTG

1420 1440 1460 1480 1500
ACAGTATATG CTTGACCAGC TCTCTGATCT TCAAGGTAA GAGCATATCT TGTCTGATGC CAACAGAGCT TTGTCAATGA AGGTATATGA TGATGTTCT
TGTCATATAC GAACCTGGTC AGAGACTAGA AGTTCATTTC CTCGTATAGA ACGAACTACG GTTGTCTCGA AACAGTTACT TCCATATACT ACTACAAAGA

1520 1540 1560 1580 1600
CTCTCTCTCC TCCAGTTTCT ATTTATAGAT GGAACCTTTA AATAGTCCAA TTTATATATA TGAGTCTAAA TTTCACTTC TTCAACTGCT ACATGTTTCT
GAGAGAGAGG AGGTCAAAGA TAAATATCTA CCTTTGAAAT TTATCAGGTT AAATATATAT ACTCAGATTT AAAGTGAAG AAGTTGACGA TGTACAAAGA

1620 1640 1660 1680 1700
TTTGTATTAT TTCTATGATA TCTTCAGGAA AGTTTGAAAA ATATTGTGTT TTGTTTAGCT GGAAGATATG ATCGGCGTGA GACATCACCA TATAGGAGGA
AAACATAATA AAGATACTAT AGAAGTCTTT TCAAACTTTT TATAACACAA AACAAATCGA CCTTCTATAC TAGCCGCACT CTGTAGTGGT ATATCTCTCT

1720 1740 1760 1780 1800
GGATGGGAAG GTGGTGATCA ACAGAATATT GCCTATGGAC ATCCTCAGGC TCATTCTCAG GGAATATACC AATCTCTTGA ATGTGATCCC ACTTTGCAAA
CCTACCTTTC CACCACTAGT TGTCTTATA CGGATACCTG TAGGAGTCCG AGTAAGAGTC CCTGATATGG TTAGAGAACT TACACTAGGG TGAACGTTT

1820 1840 1860 1880 1900
TTGGGTAAAT CAAACAACCT TTCTTGCTTT AAGACATCAA CTTAGGTTAT AAACAGTTAG CAGTTTGCTT TAAGCCCAAC ATTGCTTTTG TTTTATAGAG
AACCCTTTA GTTTGTTGAA AAGAACGGAA TTCTGTAGTT GAATCCAATA TTGTGCAATC GTCAAACGAA ATTGCGGTTG TAACAGAAAC AAAGTATCTC

1920 1940 1960 1980 2000
GCTTTGGTTA AAATCTGCTG TGTTTAGTCT AAGGATTGAG CACTTTGATG TCTGAAGTAT GGAAATCAA TATCTCAGAC TTGAAAATGT GGGTTTCTAT
CGAAACCAAT TTTGAGCACA ACAATCAGA TTCCTAAGTC GTGAACTAC AGACTTCATA CCTTTTAGTT ATAGAGTCTG AACTTTTACA CCCAAAGATA

2020 2040 2060 2080 2100
TGTTGACTTC GAAACTATGT TGTGTGGTG TTGCAACAG ATATAGCCAT CCAGTGTGCT CAGAGCAAAT GGCTGTGACG GTGCAAGGTC AGTCCCAACA
ACAACCTGAG CTTTGATACA ACAACACCAC AACGTTTGTC TATATCGGTA GGTACACAGA GTCTCGTTTA CCGACACTGC CAGTTCCAG TCAGGGTTGT

2120
AGGAAACGGC TACATCCCTG GCTGGATGCT G
TCCTTTGCCG ATGTAGGGAC CGACCTACGA C

SEQ ID NO:43

Arabidopsis SEP3 genomic sequence

-2981 -2961 -2941 -2921 -2901
GTCCCCTCC CATTACGCTC TGACGTGGAC CCTGTCGGTC TATTTTATAG AGATTAAATCC AACGGTTCTT ATTCTTTCTT CGACCCCTCA CGACATTGCC
CAGGGGAAGG GTAATGCAGA ACTGCACCTG GGACAGCGAC ATAAAAATCG TCTAATTAGG TTGCCAAGAA TAAGAAAGAA GCTGGGAAGT GCTGTAAACGG

-2881 -2861 -2841 -2821 -2801
TCAAAGCCGT CCGATTCTCA TCTCAGCCCC AATGGACCAC ATATATCACC AGTACTCCGC AACTTAGCTG TCGTGATAGG TTTCACGTGG CATTTATTTG
AGTTTCGGCA GGCTAAGAGT AGAGTGGCGG TTACTCGGTG TATATAGTGG TCATGAGGCG TTGAATCGAC AGCACATCCT AAAGTGACCC GTAAATAAAC

-2781 -2761 -2741 -2721 -2701
TTCTAGTTTG TAGTGCAAACT ATTGCAAGTT GATATGCTCC CCTATCGATC ACCGTCGTCT CTTTAGCTTC ACATCGAGAT TCTTCTTTCT TTCCTACGTG
AAGATCAAC ATCAGCTTTG TAACGTTCAA CTATACCAGG GGATAGCTAG TGGCAGCAGA GAAATCGAAG TGTAGCTCTA AGAAGAAAGA AAGGATGCAC

-2681 -2661 -2641 -2621 -2601
TAATAGCATT TTTGATTTTG AGAATTTCTT TAGAACCGTT GGATCTCTCA TCGTTGGTTG ATCCATCCAT CCAATGGA CCGTGTGTG CTCCATCCAG
ATTATCGTAA AAATAAAAC TCTTAAAGAA ATCTTGCGAA CCTAGAGAGT AGCAACCAAC TAGGTAGGTA GGTTTACCCT GGACACACAC GAGGTAGGTC

-2581 -2561 -2541 -2521 -2501
GGCATATGAT CCCAAAGCCA AAAGAGTATT TCCAAGTGCT TTCTTTCTTT CTCTCTTTCT TTCTTACTAA CCTTTTTTTT TCTTATGCTT TAGACTAAGA
CCGTATACTA GGGTTTCGGT TTTCTCATAA AGGTTACGA AAGAAAGAAA GAAAGAAAGA AAGATGATT GGAAAAAAA AGAATACGAA ATCTGATTCT

-2481 -2461 -2441 -2421 -2401
 AATTTATCG GCCATATCCA CTTTACGAA TACTTCTT ACAAGATCTA GATTTTITG AGTTAATTCG GTGTATATAA CATTGCGATG GACTGCAATT
 TTAATAAAGC CGGTATAGGT GAAAATGCTT ATATGAAGAA TGTTCAGAT CTAACAAAAAC TCAATTAAGC CACATATATT GTAACCGTAC CTGACGTTAA

-2381 -2361 -2341 -2321 -2301
 AAGTAATGGT AATGTGATCA TGATGCGATG TGTGTTATC AGTAGTATAA TATTGATGGG CTACCCTGGA AAACAAAATT ACGTGTATATA TGTACACAAT
 TTCAATACCA TTACACTAGT ACTACGCTAC ACAGCAATAG TCATCATATT ATAACCTACC GATGGGACCT TTTGTTTAA TGCACAATAT ACATGTGTTA

-2281 -2261 -2241 -2221 -2201
 TTGGTAGAAC CGTAGAAATT AAAGTGAATA AAACCTTCTA TAATGTTCAA AATTATATGG TACAGATTAA TACGAAAAA CATTACGCTT TTACGTAAAC
 AACCATCTTG GCATCTTTAA TTTGACTTAT TTTGGAAGAT ATTACAAGTT TTAATATACC ATGTCTAATT ATGCCTTTTT GTAAGTGGCA AATGCATTGT

-2181 -2161 -2141 -2121 -2101
 ATTAAGTGA AAGTAAATTT ATCCAAAAA TATTATATC ACATCAATGT TATATTCTA AGTTTITTTA TATCTCTAAT GGTATATGTT TTACAGATTG
 TAATTCACCT TTCATTTTAA TAGGGTTTTT ATAAATATAG TGTAGTAAAC ATATAAAGAT TCAAAAAAAT ATAGAGATTA CCATATACAA AATGTCTAAC

-2081 -2061 -2041 -2021 -2001
 TTTTTTGGGA AAATCTTAA AGAGACTTGA AGAATGTTTT TTTTTATTT TCTTGAATG TTTGACACTT GAAACCGTTT AAAAAGTCAA ATATAGTATA
 AAAAAACCTT TTTAAGAAAT TCTCTGAAC TCTTACAAAA AAAAATAAAA AGAAGTTTAC AAACGTGAA CTTTGGCAAA TTTTGTAGTT TATATCATAT

-1981 -1961 -1941 -1921 -1901
 TATCATGTTT GGTCTCATAC CTGTGAATTC ACCACATATA TTATCAATGG GGAAGATTTG AAAATTTTTG GGGGATCACA AAACGAAGGA AAGAGTACAA
 ATAGTAACAA CCAGAGTATG GAACATTAAG TGGTGTATAT AATAGTTACC CCTTCTAAAC TTTTAAAAAC CCCCTAGTGT TTTGCTTCCT TTCTCATGTT

-1881 -1861 -1841 -1821 -1801
 AAAGAGAAGG AAAAGATAGA AGATATATGT TTTTAACTTC ATTGGTATGA CATCAATAAA TAAATAGTTG AATGTACTTT AGTTTCTCTT TTGGTTTAAAT
 TTTCTCTTC TTTTCTATCT TCTATATACA AAAATGGAAG TAACCATACT GTAGTTATTT ATTTATCAAC TTACATGAAA TCAAGAGAAA AACCATAATTA

-1781 -1761 -1741 -1721 -1701
 GCACATCATC TCGATCAATT GTCATCATCT TACATTGAAT TATACGACCA GATCTGATAA CAAGTGAATT CGTACTTGCC CTTCCCTTTC TTCTCATACG
 CGTGTAGTAG AGCTAGTTAA CAGTAGTAGA ATGTAACTTA ATATGCTGGT CTAGACTATT GTTCACTTAA GCATGAACGG GAAGGGAAAG AAGAGTATGC

-1681 -1661 -1641 -1621 -1601
 TCCTTCTAAC TAATTTTGAT TGTAACCTAT AATTATATATA CCATATTTAA TTTTATTTTA TCTAAAACCA ATTGAAGCAA ATTAATAATAT CATAAATCTT
 AGGAAGATTG ATTAATACTA ACATTGAATA TTAATATATT GGTATAAATT AAAAATAAAT AGATTTTGGT TAACCTCGTT TAATTTTATA GTATTTAGAA

-1581 -1561 -1541 -1521 -1501
 GAGTCCACA TGAAGACAAT ATATAAACT CGTGCAAAAT TGCTTAAAT GCTTCTATGA GACCATGACC AAGTGAGATT AATAAGCGAT TCAATGTGCA
 CTCAGGGTGT ACTTCTGTTA TATATTTTGA GCACGTTTAA ACGAATTTTA CGAAGTACT CTGGTACTGG TTCCTCTAA TTATTCGCTA AGTTACACGT

-1481 -1461 -1441 -1421 -1401
 AATCAAAAGA GAAAGAAGC TAATGGGTTT AAATATAACC AAACAGAATA ATAATGCTAT GTTTAGTTTT TCTAATGAA TCATACCTTT GTGTCCATCA
 TTAGTTTTCT CTTTCTCTCG ATTACCCAAA TTTATATTGG TTTGCTTAT TATTACGATA CAAATCAAAA AGATTAACTT AGTATGGAAA CACAGGTAGT

-1381 -1361 -1341 -1321 -1301
 CCTACTTACC GGTCAAGATA AAGCAATTAC GTCTGCAACC AAAAGCACT AAGACTTTTC GTGAGACATG ATCTCTAACA TCGGACGAAC CCTAAGATAA
 GGATGAATGG CCACTCTTAT TTCGTTAATG CAGACGTTGG TTTTTCGTGA TTTTGAAAGC CAGTCTGTAC TAGAGATTGT AGCCTGCTTG GGATTTCTAT

-1281 -1261 -1241 -1221 -1201
 CCAAAATAAA CTATATCTTA TATCAAATC TCTGTTTATT TTATCCATTT ATGTTTCTTT TCTTTCCCAT AATTTTTTTT GTGTCTCATC AGACTCTCTT
 GGTTTTATTT GATATAGAAT ATAAGTTTAG AGACAAATAA AATAGGTAAA TACAAAAGAA AGAAAGGGTA TTAACAAAAA CACAGAGTAG TCTGAGAGAA

-1181 -1161 -1141 -1121 -1101
 ACCAAACTGA ATTTATCAAC ATGGTTTTT TTTTGGCCAC ATCAAAATGG TGGTTTATAA AGTAGACTAA TACAAAAGAC ATTTCTGTTA ATTTCACTAA
 TGGTTTGACT TAAATAGTTG TACCAAAAAA AAAACCGGTG TAGTTTACC ACCAAATATT TCATCTGATT ATGTTTTCTG TAAAGACAAT TAAAGTGATT

-1081 -1061 -1041 -1021 -1001
 CAAAAAATAT CTTAGCAGTA CTATAGATTG GAAAGGAAA AGCAATCTA GCAATAGAT TTATCAAAAC TAGCAGTAAG AGTTTATAGAT ATCATGAAAA
 GTTTTATTA GAATCGTCAT GATATCTAAC CTTTCTCTT TCGTTTAGAT CGTCATTCTA AATAGTTTGG ATCGTCATTCT TCAAAATCTA TAGTACTTTT

-981 -961 -941 -921 -901
 CATCACAAAC GAGTAGTGT TACTTTTACA TTTTAAACCA ATCACAAGGG TAGTCCGTA AGTTGGGAAA ATCGTACGAG GCTTACCTTA GTTAAGGTTA
 GTAGTGTGTT CTATACACAA AATGAAATGT AAAAATGGT TAGTGTTCCT ATCAAGGCAT TCAACCTTTT TAGCATGCTC CGAAGTGGAT CAATTCCAAT

-881 -861 -841 -821 -801
 GGTCACATGA TTCCCTGAAC TCGATTTTAT AAGTAAAAA GAAAAATTTA TAAATCAAAA ATTTTTTATA TAAAAAATC AGGTGGATT ATCAGACCTC
 CCAGTGACT AAGGACTTG AGCTAAAAA TTCATTTTTT CTTTTTAAAT ATTTTAGTTT TAAAAAATAT ATTTTTTTAG TCCACCTAAA TAGTCTGGGA

-781 -761 -741 -721 -701
 ACCATCGAGA TGTGACACG TGTCCAACT CATTCAATGC CTTACTATTT TCTGTTTAGG GTTGCAATCA CTCATCGCAC ACGCGCCATC TCCACCTTCC
 TGGTAGCTCT ACAGCTGTGC ACAGGTTTGA GTAAGTAACG GGATGATAAA AGACAAATCC CAACGTTAGT GAGTAGCGTG TGCSCGGTAG AGGTGGAAGG

-681 -661 -641 -621 -601
 ATTATTAATC TCTCATTTTC AACATCACAC TCTTACGAAT CATACGATTT TAATATCTCT GTCTCTCTCA ACGTATTAAA TAAAAATGGT TTTAAATGTT
 TAATAATTAG AGAGTAAAG TTGTAGTGTG AGAATGCTTA GTATGCTAAA ATTATAGAGA CAGAGAGAGT TGCATAATTT ATTTTACCA AAATTTACAA

-581 -561 -541 -521 -501
 AGGGTTTTTT GTAGGATTTT CAATTATTAA TCTCTATAAT TCGATGAAC AAGTAAAAA GCATCAAACT TTCTTGGCAG ATCAGATTTT TCTCTAAACT
 TCCCAAAAAA CATCTAAAA GTTAATAATT AGAGATATTA AGCTACTTGA TTCATTTTTT CGTAGTTTGA AAGAACCCTC TAGTGTAATA AGAGATTGTA

-481 -461 -441 -421 -401
 AAATATGGAC TGAATTGAA AAATTAACCT ACTAGCTAGA ATAAAGTGT GGTGAGAGTG GAACCTTAAT TTCTCTCCTT TACTAATTAT GTATAAACAC
 TTTATACCTG ACTTTAACTT TTTAATTGTG TGATCGATCT TATTTACAA CCACCTCTCAC CTTGAGATTA AAGAGAGGAA ATGATTAATA CATATTTGTG

-381 -361 -341 -321 -301
 AAAAAATGAC CAAATTTTAA GGTTTGAAA TATCTAAGCA TGGATAGGGT AATTAACATT TTTTCTTTCA ATTTTGCAAT ATTTGAATAA ATCCTATGAG
 TTTTACCTG GTTTAAAAAT CCAAACTTTT ATAGATTCGT ACCTATCCCA TTAATTGTAA AAAAGAAAGT TAAACGTTA TAACTTATT TAGGATACTC

-281 -261 -241 -221 -201
 GGTCTTTGGT ACACAATAAT TGGAGGGTAT ATAGTTGAGT CTGAGAGTAT ATTAGAAAGA GAATATTTCA AGTAATGAAG CTGACATGTT TATATGTACT
 CCAGAAACCA TGTGTTATTA ACCTCCCATTA TATCAACTCA GACTCTCATA TAATCTTTCT CTTATAAAGT TCATTACTTC GACTGTACAA ATATACATGA

-181 -161 -141 -121 -101
 TTGAGAGAAG TGTGTGAGA TTTGTACAAA TGTATATGTA CACTTTAAAA AGCAATATAA GATAGATAAA AAAAAATATA AGAAAAAAG AAAGAAAGAA
 AACTCTCTTC ACAACACTCT AAACATGTTT ACATATACAT GTGAAATTTT TCGTTATATT CTATCTAATT TTTTATATT TCTTTTCTC TTTCTTTCTT

-81 -61 -41 -21 -1
 AGAAAGAAAG AGAGAGGCTC ATATATATAT AGAATTGCTT GCAAGGAAAG AGAGAGAGAG AGATTGAGAT ATCTTTTGGG AGAGGAGAAA GAAAAAGAAA
 TCTTTCTTTC TCTCTCCGAG TATATATATA TCTTAACGAA CGTCTCTTC TCTCTCTC TCTAACTCTA TAGAAAAACC TCTCTCTTCT CTTTCTTCTT

20 40 60 80 100
 ATGGGAAGAG GGAGAGTAGA ATTGAAGAGG ATAGAGAACA AGATCAATAG GCAAGTGACG TTTGCAAAGA GAAGGAATGG TCTTTTGAAG AAAGCATACG
 TACCCTCTCT CCTCTCATCT TAACTTCTCC TATCTCTTGT TCTAGTTATC CGTTCACGTC AAACGTTTCT CTTCCTTACC AGAAAACTTC TTTTGTATGC

120 140 160 180 200
 AGCTTTTCACT TCTATGTGAT GCAGAAGTTG CTCTCATCAT CTCTCAAT AGAGGAAAGC TGTACGAGTT TTGCAGTAGT TCGAGGTATA TATCTACTTT
 TCGAAAGTCA AGATACACTA CGTCTTCAAC GAGAGTAGTA GAAGAGTTTA TCTCTTTTCC ACATGCTCAA AACGTCATCA AGCTCCATAT ATAGATGAAA

220 240 260 280 300
 TGTATATATA TTAATTTATA CATAAACATT TTATATACAT ATTAAGTAAC AAAAAAATGT CTTGTATGTA TGGGTCTCTC TGTGATGTGT TGTGTGTGCG
 ACATATATAT AATGAATATT GTATTGTGTA AATATATGTA TAATTCATTG TGTTTTACA GAACATACAT ACCCAGAGAG ACACCTACACA ACAACACAGC

320 340 360 380 400
 TACGTACGTC TTCTATCATA TCTTTTAAA AGAAGCAAAG AGGAAAAAAA ATTTGGGATA CCCCAAATCT GTATCATTTT ATAACAAGTT TGCTTTTGTG
 ATGCTATGAC AAGATAGTAT AGGAAAAATT TCTTCGTTTC TCCTTTTTTT TAAACCTTAT GGGGTTTAGA CATAGTAAAA TATTGTTCAC ACGAAAAAAC

420 440 460 480 500
 ATGTTCTTTT GTGTTTCTCT TTGATTTCOA TTTTGTGTTT TGATTTTCTT TCTATTTCTC TTTACATCTA TCAAAGTTT TTTTCTTATA TTTTATTGCT
 TACAAGAAAA CACAAGAGA AACTAAAGGT AAAAACAAAA ACTAAAAAAA AGATAAAGAG AAATGTAGAT AGTTTCAAAA AAAAGAATAT AAAATAACGA

520 540 560 580 600
 TATTGTGTTG TCTACTTAAT TCACATTATC TGAGAGAAGA ACAATCTATC TGATATGAAA TTAGGGTTAA TTTCTCTGTT GAGTACTCTT TAATTCACAT
 ATAAACAAAC AGATGAATTA AGTGAATAG ACTCTCTTCT TGTTAGATAG ACTATACTTT AATCCCAATT AAAGAGAACA CTCATGAGAA ATTAAGTGTGTA

620 640 660 680 700
 AAGCTTAAAG TTTCACCTT TTGATTCTGG GGGTCGTCCA ATTCGATCAA ATCACTCAAT TTTGTTGTCA GATTGATATA AGTTTATAGG GGGATATTGT
 TTCGAATTTT AAAGGTGGA AACTAAGACC CCCAGCAGGT TAAGCTAGTT TAGTGAGTTA AAACAACAGT CTAATATAT TCAAGTATCC CCCTATAACA

720 740 760 780 800
 TTCCACGACA ATCCATTTTA GTAACCTTAA GGGGTTTCCA ATTTTGGGTT TTGAATTGAC GCTAATGTCA AATTCATCTA AAGTCCGTTG GATATGTATA
 AAGGTGCTGT TAGGTAAAA CATTGGGAAT CCCCAGAGGT TAAACCCCAA AACTTAACTG CGATTACAGT TTAAGTAGAT TTCAGGCAAC CTATACATAT

820 840 860 880 900
 CTTGGGGATG GGATTCATCC TTTTCTCTGG GTTCTTTAGA TCTTCTCTTA AAAGACTAAC AGATTTTGTT GTAAACCCTA GGAAACAGTT AAAAATCCCA
 GAACCCCTAC CCTAAGTAGG AAAAAAGACC CAAGAAATCT AGAAGAGAAT TTTCTGATTG TCTAAAAACA CATTGCGGAT CCTTTGTCAA TTTTTAGGGT

	920	940	960	980	1000
	TTTTTAAAAA	CATGTTTGA	ACTTGATGAG	TAAGATTAAT	GGAAGAAATG
	ATGTTTGT	GTGGTGTGAA	GCATGCTTCG	GACACTGGAG	AGGTACCAAA
5	AAAAATTTT	GTACAAAAC	TGAACTACTC	ATTCTAATTA	CCTTCTTTAC
	TACAAAACA	CACCACACTT	CGTAGCAAGC	CTGTGACCTC	TCCATGGTTT
	1020	1040	1060	1080	1100
	AGTGTAAC	TGGAGACCA	GAACCCAATG	TGCCTTCAAG	AGAGGCGCTA
	GCAGTTGTAC	CCAATTCTCT	TCTCTTCTT	CTAATTACCT	TAATTAATTA
	TCACATTGAT	ACCTCGTGGT	CTTGGGTTC	ACGGAAGTTC	TCTCCGGAAT
	CGTCAACATG	GGTTAAGAGA	AGAGAAAGAA	GATTAAATGA	ATTAATTAAT
10		1120	1140	1160	1180
		CTCTCAATTT	TACTTTTGAT	TTTTAGAGTC	AAATGATTAA
		TGTTATAATT	TGTCATATAC	TTCAGGAAC	TAGTAGCCAG
		CAGGAGTATC	TCAAGCTTAA	GAGAGTTAAA	AATGAACTA
		AAAATCTCAG	TTTACTAATT	ACAATATTAA	ACAGTATATG
		AAGTCTTGA	ATCATCGGTC	GTCCCTCATAG	AGTTCTGAAT
		1220	1240	1260	1280
		GGAGCGTTAT	GACGCTTAC	AGAGAACCCA	AAGGTAAACT
		AATTAGCTTC	TTCAGCTACC	TTCAGAGAGT	GTTTGTTTTT
		TTAGTAGATT	TTTTTGATGG	CCTCGCAATA	CTGCGGAATG
		TCTCTTGGGT	TTCCATTGTA	TTAATCGAAG	AAGTCGATGG
		AAGTCTCTCA	CAAAACAAAA	AATCATCTAA	AAAAACTACC
		1320	1340	1360	1380
		TTTGTGATTT	GAAATAGGAA	TCTGTTGGGA	GAAGATCTTG
		GACCTCTAAG	TACAAAGGAG	CTTGAGTCAC	TTGAGAGACA
		GCTTGATTCT	TCCTTGAAGC	AAAACTACAA	CTTTATCCTT
		AGACAACCTC	CTTCTAGAAC	CTGGAGATTC	ATGTTTCCTC
		GAACTCAGTG	AATCTCTCTG	CGAACTAAGA	AGGAACCTCG
		1420	1440	1460	1480
		AGATCAGAGC	TCTCAGGGTA	CTACTTTGTT	CATCAATATC
		TTTATACACT	GATCTATTTC	CATAGTAAGA	TTAAATTG
		TGTTTAATTC	TGCAGACACA	TCTAGTCTCG	AGAGTCCCAT
		GATGAAACAA	GTAGTTATAG	AAATATGTGA	CTAGATAAAG
		GTATCATTCT	AATTTAAACC	ACAAATTAAG	ACGTCTGTGT
25		1520	1540	1560	1580
		GTTTATGCTT	GACCAGCTCA	ACGATCTTCA	GAGTAAGGTA
		AATAAAGAAA	CACCTATTCT	CCTCTCTAAA	TTCTCTCATCT
		AAAAGTAATG	TAACCAAGAA	CAAAATACGAA	CTGGTCGAGT
		TGCTAGAAGT	CTCATTTCCAT	TTATTTCTTT	GTGAGTAAGA
		GGAGAGATT	AAGGAGTAGA	TTTTCATTAC	ATTGGTTCTT
		1620	1640	1660	1680
		AACACAAATA	TTTGAGGCAG	GAACGCATGC	TGACTGAGAC
		AAATAAACT	CTAAGACTAA	GGGTAATTAA	TATACATTCT
		CATATCACCA	AATTAATGCA	TTGTGTTTAT	AAACTCTGTC
		CTTGCGTACG	ACTGACTCTG	TTTATTTTGA	GATTCTGATT
		CCCATTAAIT	ATATGTAAGA	GTATAGTGGT	TTAATTACGT
		1720	1740	1760	1780
		TCACTAAATT	TGGTTTAAAT	GTGTGTGTGT	ATATACATAT
		GTGACAGTTA	GCTGATGGGT	ATCAGATGCC	ACTCCAGCTG
		AACCCTAACC	AAGAAGAGGT	AGTGATTAA	ACCAATATTA
		CACACACACA	TATATGTATA	CAGTGTCAAT	CGACTACCCA
		TAGTCTACGG	TGAGGTGCGAC	TTGGGATTTG	TTCTCTCCCA
		1820	1840	1860	1880
		TGATCACTAC	GGTCGTATC	ATCATCAACA	ACAACAACAC
		TCCCAAGCTT	TCTTCCAGCC	TTTGAATGT	GAACCCATTC
		TTCAGATCGG	GTAACCTTAG	ACTAGTGATG	CCAGCAGTAG
		TAGTAGTTGT	TGTTGTGTGT	AGGGTTCGAA	AGAAGGTGCG
		AAACCTTACA	CTTGGGTAA	AAGTCTAGCC	CATTGAAATC
		1920	1940	1960	1980
		ACTAGTATAA	CCAATTGAT	TTGAGTTCTA	TTATAAGCTT
		TTCTTAAGAA	AGTATCTCAA	ACTACTAAAT	TTTATGGAGC
		AGGTATCAGG	GGCAACAAGA	TGATCATATT	GGTTAAACTA
		AATCAAGAT	AATATTGAA	AAGAATTTCT	TCATAGAGTT
		TGATGATTTA	AAATACCTCG	TCCATAGTCC	CCGTGTGTCT
		2020	2040	2060	
		TGGAATGGGA	GCAGGACCAA	GTGTGAATAA	TTACATGTTG
		GGTTGGTTAC	CTTATGACAC	CAACTCTATT	
		ACCTTACCCT	CGTCTCGGTT	CACACTTATT	AATGTACAAC
		CCAACCAATG	GAATACTGTG	GTTGAGATAA	

SEQ ID NO:44

Arabidopsis AGL20 genomic sequence

	-2981	-2961	-2941	-2921	-2901
	GAAAAA	ACACCTAAAG	AAGTGAATAT	AATAGGCATA	TACATATGAG
	CTTTTTTTT	TGTGGATTTC	TTCACTTATA	TTATCCGTAT	ATGTATACTC
					CTTTTACTTT
					TGTTTTCTC
					GCTTTTTATC
					TAAATTGGAT
					TTTCTCCTTC
		-2881	-2861	-2841	-2821
		TAAAGAGGTT	ATAAGAGGTA	AGAAAAGTAG	GACCATATAA
		TAGCTATATT	GTAGAATTTT	ATTATTGGA	GATATGGCAA
		TTTGTGTGAG	GGTCCCATGA	TTTCTCCATG	CCAGGTACT
		ATTCTCCAA	TATTCTCCAT	TCTTTTCATC	CTGGTATATT
					ATCGATATAA
					CATCTTAAAA
					TAATAAACCT
					CTATACCGTT
					AAAAACACTC
					CCAGGTACT
		-2781	-2761	-2741	-2721
		AGACTAAAGT	GTGAGCAGC	ATTTATCTTT	GTAATTAATA
		AAATAATAAA	TATATTATTA	TTGTCTCGGG	ATTTTTCGAT
		TGATGAGAAA	AAGTAAGAGG	TCTGATTTCA	CACCTCGTGC
		TAAATAGAAA	CATTAAATTAT	TTTATTATTT	ATATAATAAT
					AACAGAGCCC
					TAAAAAGCTA
					ACTACTCTTT
					TTCAITCTCC
		-2681	-2661	-2641	-2621
		TGCGTTTTCG	AAITATCATT	GGCTAACGTT	TGTACGTGAC
		ACGCAAAAGC	TTAATAGTAA	CCGATTGCAA	ACATGCACTG
					ACATGCCTGC
					TGCAACTACA
					TAAAGATTAT
					AACATGAGAA
					AAAAGGGTGG
					GAATAAAGAG
		-2581	-2561	-2541	-2521
		TAATCTTGT	ACATTAAACC	CAAACTAATT	TTACAACAC
		ATTGAAGACA	TGTAATTGGG	GTTTGATTAA	AATGTTTG
					TAACCACAAA
					TTAGTAACAC
					TTTAAAACTA
					AATAGATT
					ATGTGAAATA
					TACAATACTA
		-2481	-2461	-2441	-2421
		TTTGCATGAG	CTTATGACTG	GTAAACTCAT	GAGATTTC
		TATCAACATG	TTGGAAGTTA	CTAACCATAC	ATCTTTTAAA
		TGCAAAATTA	CATCATTCCT	AAACGTACTC	GAATACTGAC
		CATTTGAGTA	CTCTAAAGGT	ATAGTGTGAC	AACCTTCAAT
					GATTGATG
					TAGAAAATTT
					ACGTTTAAGT
					GTAGTAAGGA

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100

AGACTGCTAG TCTGACGATC	-2381 ACAGACATGT TGTCTGTACA	ACTTACTTAT TGAATGAATA	-2361 ACAAGGTTTT TGTTCACAAA	TCTAATCTAA AGATTAGATT	-2341 TGGCAACAAA ACCGTTGTTT	GAAACTTGTG CTTTGAACAC	-2321 ACTAARACGCA TGATTGTGCGT	TACGTATCTC ATGCATAGAG	-2301 TCATATAGTG AGTATATCAC
TAGACTAGAA ATCTGATCTT	-2281 CTCTACGTAT GAGATGCATA	CTCTCATATA GAGAGTATAT	-2261 GTCAATATTT CAGTATAAAA	TAAAAAATT ATTTTTTTAA	-2241 ATACTTTGGG TATGAAACCC	ATCTCGAAGC TAGAGCTTCG	-2221 GAAAAATTAGA CTTTTAATCT	TTAGTTTATA AATCAAAATAT	-2201 TGATTATGTA ACTAATACAT
CAAAAAAAT GTTTTTTTTA	-2181 CGGATATTAC GCCTATAATG	TACCACTTAA ATGGTGAATT	-2161 AAATAATTGT TTTATTAACA	AGTGGTCAAT TCACCAGTTA	-2141 CATATCTAAA GTATAGATTT	ATTAATCGCA TAATTAGCGT	-2121 GTGAACAAAA CACTTGTTTT	ACCTGAAGCA TGGACTTCGT	-2101 TAGCCTGGTT ATCGGACCAA
CTATCTTACT GATAGAATGA	-2081 TTCGATGTGA AAGCTACACT	CACATTACTA GTGTAATGAT	-2061 ACACGATTGT TGTGCTAACA	TTTAATCTAT AAATTAGATA	-2041 AGGACGAATC TCCTGCTTAG	CTTTAAGTAA GAAATTCATT	-2021 TGTATAGTTG ACATATCAAC	GTTCAGTTAC CAAGTCAATG	-2001 GTTAGATACAT CAATCTATGA
TTTTGTTTTG AAAACAAAAC	-1981 GATTTGTCTC CTAAACAGAG	AACCAGTTAA TTGGTCAATT	-1961 GAAGTGATCG CTTCACTAGC	TATTTACTAG ATAAATGATC	-1941 TGGTATACGA ACCATATGCT	TGATGTTTTCT ACTACAAAGA	-1921 TTAAATCTGA AATTTAGACT	ATTGGGTCTA TAACCCAGAT	-1901 CAAAATACAT GTTTTATGTA
AACTAAACTT TTGATTTGAA	-1881 CAAAACCGGT GTTTGGCCCA	TTATACTTTA AATATGAAAT	-1861 TACAAACACG ATGTTTGTGC	AAAATATAAA TTTTATATTT	-1841 GATAGAGACA CTATCTCTGT	ATTCACCAGA TAAGTGGTCT	-1821 GAAGATGTGT CTTCTACACA	ATTTATATAA TAAATATATT	-1801 AAATTTATCCA TTTAATAGGT
TACAGATTTT ATGTCTAAAA	-1781 CGGACCTATC GCCTGGATAG	TGTTTGATAT ACAAACTATA	-1761 TTAATATATA AATTTATGCA	TAAATACGTT ATTTATGCAA	-1741 AACATATTTT TTGTATAAAG	ACCAGAGAAG TGGTCTCTTC	-1721 ATGTGTATTT TACACATAAA	TTCGAAATAA AAGCTTTATT	-1701 TTAGTTTGTG AATCAAAACAC
TGGTCTCTCT ACCAGGAGGA	-1681 CCCAGATATG GGGCTATATC	ATAAAAGATC TATTTTCTAG	-1661 ATTAGATATC TAATCTATAG	GATTAACAAT CTAAATTGTTA	-1641 TTTATCTCCA AAATAGAGGT	AAAAAGGATA TTTTCCTCAT	-1621 TTTTTTTGGT AAAAAAACCA	GCCACTAGCT CGGTGATCGA	-1601 AGACAAGACG TCTGTTCTCG
TTCGATAAGC AAGCTATTTC	-1581 TGAATTATTA ACTTAATAAT	TTGGATTCTT AACCTAAAGA	-1561 AAGTTACGTT TTCAATGCAA	TTCTTTAGTA AAGAAATCAT	-1541 ATCCGAGGGA TAGGCTCCCT	CCAAAAATAG GGTTTTTATC	-1521 CAAAATGCCTC GTTTACGGAG	TTTATAGACG AAATCTGTGC	-1501 TCGCTACTTA AGCGATGAAT
ACGCCATTGC TGCGGTAACG	-1481 CCCATTGTCT GGGTAACAGA	CTGTACTAGC GACATGATCG	-1461 CTCCAAATAT GAGGTTTATA	TTGGATTAAT AACCTAATTA	-1441 GGTCACTTAG CCAGTGAATC	GTAATGAGGA CATTACTCTC	-1421 AATTGTAGTA TTAACATCAT	TTTTGTAATG AAAAATTAC	-1401 TGGTTTTGTC ACCAAAACAG
CAACTTATAA GTTGAATATT	-1381 AAACTTACAA TTTGAATGTT	TTGCAAGTAA AACGTTCAAT	-1361 TTAATTATTC AATTAATAAG	ACATGGAGAT TGTACCTCTA	-1341 GTAAGATTAT CATTCTAATA	GTCATATAAC CAGTATAATTG	-1321 TAAAAACACA ATTTTGTGTT	ATTTAAGAAC TAAATCTTGT	-1301 AACAATAAGA TTGTTATTTCT
AACAATGACG TTGTTACCTG	-1281 AAACAAGCAT TTTGTTCGTA	AGAAAAATATA TCTTTTATAT	-1261 CAAATCAAAAT GTTTAGTTTA	GAATTTTATC CTTAAAAATAG	-1241 TGTGTGGATG ACAACCTTAC	GAAAGATATT CTTTCTATAA	-1221 ATAAAAATTG TATTTTAAAC	ATTAATAACCA TAATTTTGGT	-1201 ATATAGTTGT TATATCAACA
ATTACTCACA TAATGAGTGT	-1181 GGTAAGAAAA CCATCTTTTT	AACGATATTC TTGCTATAAG	-1161 TTATTTTTC AATAAAAAATG	TATCAATTAC ATAGTTAATG	-1141 AAGTGGGGGC TTCACCCCGC	ATATAGGTAC TATATCCATG	-1121 GAGAGAGTGT CTCTCTACAC	TTCTGTCCAC AACACAGGTG	-1101 ATTAATAACA TAATTTTTGT
AAAAAAGATT TTTTTTCTAA	-1081 TTTGTTAGAA AAACAATCTT	GAAATTTAAT CTTTAAATTA	-1061 AAAAATAATT TTTTTATTA	TGACAGGCAT ACTGTCCGTA	-1041 TTCCATCCAA AAGGTAGGTT	CTAGATATTT GATCTATAAA	-1021 ATGGGAGGTA TACCCTCCCT	AAAAGATGTG TTTTCTACAC	-1001 TATGTAAAAA ATACATTTTT
TGTCCATATG ACAGGTATAC	-981 TATCAAAATA ATAGTTTTAT	TGCTATTTTT ACGATAAAAA	-961 GGTCTTTCTT CCAGAAAGAA	AAGGCTTTTT TTCCGAAAAA	-941 TCCAAAAATA AGGTTTTATT	GTAAGAGGAT CATTTCTCTA	-921 AGGTTTCAAG TCCAAAGTTC	CGTCCATCAT GCAGGTAGTA	-901 ATTTGCGACA TAAACGCTGT
CATATGACTG GTATACTGAC	-881 ACTATTTAGC TGATAAATCG	TCCCTCCCTC AGGAGGGAGA	-861 TTCTTTCTCT AAGAAAGAGA	TATTTTATTA ATAAAATAAT	-841 TCCTTCTCCA AGAAAGAGGT	AGAAATAAAA TCTTTTTTTT	-821 TAGAAAAGAA ATCTTTTCTT	AATATATATG TTATATATAC	-801 GTTTCACAAA CAAAGTGT
CACCATTACC GTGGTAATGG	-781 ATAACTACAA TATTGATGTT	CGAGAAGAGG GCTCTTCTCC	-761 ATCTTTTTTT TAGAAAAAAT	AGGAGAAAAA TCCTCTTTTC	-741 CAGAGAGAGA GTCTCTCTCT	AGAGACGAGT TCTCTGCTCA	-721 GTGTGAAGTT CACACTTCAA	TTTTTGTCTT AAAAACAGAA	-701 TTGTTTCTTT AACAAGAGAA
TATTACACAC ATAATGTGTG	-681 AAATAGATGA TTTATCTACT	AACGAGGAAA TTGCTCTCTT	-661 GCTACTTCTT CGATGAAGAA	TTGCTACTTC AACGATGAAG	-641 CATAAAAAAG GTATTTTTTC	TTCTTCTCTT AAGAAGGAAA	-621 CGCAGAGAGT GCGTCTCTTA	CAACTTTGAT GTTGAAACTA	-601 CATCTTCTTC GTAGAAGAAG
CTTCTCTTTC GAAGAGAAAG	-581 TTTCTTCTTC AAAGAAGAAG	TCCCTCCAGT AGGGAGGTCA	-561 AATGCTTATA TTACGAATAT	TAGTCTCTCT ATCAGAGGAG	-541 CTATATCTCT GATATAGAGA	ACCTATACAT TGGATATGTA	-521 ACACAAACCC TGTTTGGG	TTTATCTCTG AAATAGGAGC	-501 AAAGCTTCTC TTTCGAAGGA
CCTGTTTAGG GGACCAATCC	-481 TTTTTATCAA AAAAATAGTT	ACCCTTTTAG TGGGAAAAAT	-461 CCAATCGGTA GGTTAGCCAT	AGATCTCTTC TCTAGAGAAG	-441 GTCATGATCT CAGTACTAGA	TTTCTTTTTT AAAGAAAAAA	-421 CTTTTGTCTT GAAAAACGAA	GTACTCTGAT CATGAGACTA	-401 GGATCTATAA CCTAGATATT
ACTTATATGG TGAATATACC	-381 GTTTGGTTTC CAAACCAAAG	ATTTGGTTTC TAAACCAAGC	-361 ATTTGATGTG TAAACTACAC	TTTGGTTTCT AAACCAAAGA	-341 TTGTCCTAAA AACAGGATTT	TCTCATGAAA AGAGTACTTT	-321 GGAGGTTGCA CCTCCAACGT	TCCTTCAATT AGGAAGTTAA	-301 AAACCGATAA TTTGGCTATT
CAAAAGTTTC GTTTCAAAAG	-281 CATTACAGAC GTAATGTCTG	TTATAGATCA AATATCTAGT	-261 GATACTTTAG CTATGAAATC	ATTGTTTTGC TAACAAAACG	-241 TTTTTGGGTA AAAAACCCAT	CTTAATCTTT GAAATTAGAA	-221 CGTTGACTTC GCAACTGAAG	ATCAGTCTTC TAGTCAGAAG	-201 TCCCACCCAA AGGGTGGGTT
ACAAAAAAGT TGTTTTTTCA	-181 CATATTTTGA GTATAAAGCT	TCATATCTTC AGTATAGAAG	-161 ATTTTTTTAA TAAAAAAATT	CCTACTCTCT GGATGAGAGA	-141 TTGATTTCATA AACTAAGTAT	TATGAAATGG ATACTTTTACC	-121 GTTGTTTTAT CAACAAAATA	GTGTGTGACT CACACACTGA	-101 AATCTTGTGA TTAGAACAAT
TTGAGGTGGT AACTCCACCA	-81 TGCAACATTG ACGTGGTAAAC	ATCTACCGTT TAGATGGCAA	-61 TTCTTCAATT AAGAAGTTAA	TTTGAaaaaa AAACTTTTTT	-41 TAATTTTATT ATTAAAAATA	TTTTTTCTGT AAAAAAGACA	-21 GTGCAAGGGA CACGTTCCCT	AATTAACATA TTAATTGATT	-1 AGAAGAAGAT TCTTCTCTCA
ATGGTGAGGG TACCACCTCC	20 GCAAACTCA CGTTTTGAGT	40 GATGAAGAGA CTACTTCTCT	60 ATAGAGAATG TATCTCTTAC	80 CAACAAGCAG GTTGTCTGTC	100 ACAAGTGACT TGTTCACTGA	120 TTCTCCAAAA AAGAGGTTTT	140 GAAGGAATGG CTTCCCTACC	160 TTTGTGAAG AAACAACCTC	180 AAAGCCTTTG TTTCGGAAAC
AGCTCTCAGT GCTTTGTGAT	120 GCTGAAGTTT	140 CTCTTATCAT	160 CTTCTCTCCT	180 AAAGGCAAC	200 TTTATGAATT	220 CGCCAGCTCC	240 AAGTACGTTT	260 TTTTTGTCTT	280 TTTTTGTCTT

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90

```
TCGAGAGTCA CGAAACACTA CGACTTCAAA GAGAATAGTA GAAGAGAGGA TTTCCGTTTG AAATACTTAA GCGGTCGAGG TTCATGCAAG AAAAAACAGAA
220 240 260 280 300
TCTTACAAAT CATCCATAGA AAGAGAGAGA GAGAGAGATC TCATTAACCT CTCTATTGTG ATCTTAATTT TTTTGGTTT ATATATGGAT TTGATTGGCC
AGAATGTTTA GTAGGTATCT TTCTCTCTCT CTCTCTCTAG AGTAATTGGA GAGATAAAAC AAAAACCAAA TATATACCTA AACTAAACCGG
320 340 360 380 400
TTTGTGGGAA TCACATCTCT TTGACGTTTG CTTTGAGAGG TGTGTTTAAA TGAGTTTCTT GGTTCCTGCA AAATTAGGGC TATTATTAAA GTAGTATCAA
AAAAACACCT AGTGATAGAA AACTGCAAAAC GAAACTCTCC ACACAAATTT ACTCAAGAAA CCAAGAGACGT TTTAATCCCG ATAAATAATT CATCATAGTT
420 440 460 480 500
GTACATATAC CCTCTATTCT ATTGTTTTTT TATTTCCGCT AGTATATCAT CTTGTTTAAAT CATCTGTCTC TCTCTTTCTC AATTAGTTTC TCAAGTTATG
CATGTATATG GGAGAATAAA TAACAAAAAA ATAAAGGCGA TCATATAGTA GAACAAATTA GTAGACAGAG AGAGAAAGAG TTAATCAAAAG AGTTCAATAC
520 540 560 580 600
ATATAAAATA AATGTGCTCT TTCGTAGCCA ATTTACACTT GTTATATATT TGATCTTCTT AGAGATCATG ATCACAATAGT ATTAATAAAA CAACTTTCAA
TATATTTATT TTACACGAGA AAGCATCGGT TAAATGTGAA CAATATATAA ACTAGAAGAA TCTCTAGTAC TAGTGTATCA TAATTATTTT GTTGAAAGTT
620 640 660 680 700
TTAGTATTCT TTTGGTTTGA ACTAATCTTT GTCTTGTATT TGCTTTAAGC AAAACATGTT GTTCTAAATT CTAAGTGATG ATTAGGAAGT TGTTCATCA
AATCATAAGA AAACCAAACT TGATTAGAAA CAGAACAATA ACGAAATTCG TTTTGTACAA CAAGATTAAA GATTCACTAC TAATCCTTCA ACAAGTAGT
720 740 760 780 800
TTCTGTGATT ATTAATCCCT CATGCTTCAT TTCATGCTCA TTCTTAATTT AGTTCAATTT GTTTGAATAT TTGTTCTCGA TTTTGACATA GAAACTCAAA
AAGGACTAAA TAATTAGGGA GTACGAAGTA AAGTAGAGT AAGGATTAAA TCAAGTTAAA CAAACTTATA AACAAAGGAT AAAACTGTAT CTTTGAAGTT
820 840 860 880 900
GCTAGCTAGC CAAACCTAAA TGTGTGATTG TTTTGAGAAAT CAAAAGAGTT TTATCTTGTA CTGTTAGGTA GTAGGGAAGC CAAACTTACT TTTGATGAAT
CGATCGATCG GTTTGGATTT ACAACTAACA AAAACTCTTA GTTTTCTCAA AATAGAACAT GACAATCCAT CATCCCTTTG GTTTGAATGA AAACCTACTA
920 940 960 980 1000
CATTACTTCT GTAATGAAA ATGCCAGCTT TTGATCAGAT GTTTCAGACA TTTGGTCCAT TTGGGAAAGT ACTTCTTTCT CTCGAACCTA CTAAATATA
GTAATGAAGA CATTTACTTT TACGGTCGAA AACTAGTCTA CAAAGTCTGT AAACAGGTA AACCTTTCA TGAAGAAGA GAGCTTGGAT GATTATATT
1020 1040 1060 1080 1100
AGATAAGACC TCACATGTTT TTGATTTTCT AAAATAGGGG GAAAAGTAC AAGACTTTTC AAGCTATGTC CTTGATTAAG TCTAGTGATA TCTTCAATA
TCTATPCTGG AGTGACAAA AACTAAAAGA TTTTATCCCC CTTTTTCATG TTCTGAAAAG TTCGATACAG GAACATAATC AGATCACTAT AGAAGTTATT
1120 1140 1160 1180 1200
GAAATGTTTT GAGAACACCA TTGGGATCTA AATTGATCT CTGATGATT ACTTTAATGT TCCAATTATA TATGTTTTTG ACAGTATGCA AGATACCATA
CTTTACAAA CTCTTGTTGGT AACCTAGAT TTAACCTAGA GACTACTAAA TGAAATTACA AGGTTAATAT ATACAAAAC TGTCATACGT TCTATGTAT
1220 1240 1260 1280 1300
GATCGTTATC TGAGGCATAC TAAGGATCGA GTCAGCACCA AACCCTTTTC TGAAGAAAAT ATGCAGSTTT ATTCTTTATG ATCTTCTTGC CTATATATCA
CTAGCAATAG ACTCCGTATG ATTCTAGCT CAGTCGTGTT TTGGCCAAAG ACTTCTTTTA TACGTCCAAA TAAGAAATAC TAGAAGAACG GATATATAGT
1320 1340 1360 1380 1400
ATTCTTGCTA ATTAATACTT TTAATATATA ATATCAAGA GCGGTAATGA ATATAACCA AATATGTATA TAATCTCAAG GTCACAGGAT CAAGTCACAT
TAAGAACGAT TAATTATGAA AATGATATAT TATAGTTTCT CGCCATTACT TATATTGGTG TTATACATAT ATTAGAGTTT CAGTGTCTCA GTTCAAGTGA
1420 1440 1460 1480 1500
ATTTATAATT AGGATATATA TGTACATGCA ATAACATTTT TGTGATATA CCAACAGCAT TTGAAATATG AAGCAGCATA CATGATGAAG AAAATTGAAC
TAAATATTAA TCCTATATAT ACAATGACGT TATTGTAAAG ACACATATAT GGTGTGCTGA AACTTTATAC TTGCTGTCTT GTACTACTTC TTTTAACTTG
1520 1540 1560 1580 1600
AACTCGAAGC TTCTAAACGG TTTGTGATAT ATACATATAT ACAAACACAT TATTCATCAC TTGTATATAT CTATTTTCATG ATGCATAGGA GAGTTTGATC
TTGAGCTTCG AAGATTGGCC AAACACTATA TATGTATATA TGTGTGTTGA ATAAGTAGTG AACATATATA GATAAAGTAT TACGTATCCT CTCGAACTAG
1620 1640 1660 1680 1700
AATTAGTGTG TTGTTTTTGT AATCAGTAAA CTCTTGGGAG AAGGCAATAG AACATGCTCA ATCGAGGAGC TGCAACAGAT TGAGCAACAG CTTGAGAAAA
TTAATCACAA AACAAAAACA TTAGTCAATT GAGAACCCCT TTCCGTATCC TTGTACGAGT TAGCTCTCTG ACGTTGTCTA ACTCGTTGTC GAACTCTTTT
1720 1740 1760 1780 1800
GTGTCAAATG TATTCGAGCA AGAAAGGTAT GTGTATATAT TTATCTGTTA TATCTCCACA TTATAAGTAT TGTTCGAATC ATCTTCTGAA ACCACTCATA
CACAGTTTAC ATAAAGCTCGT TCTTTCCATA CACATATATA AATAGACAAT ATAGAGGTGT AATATTCTA ACAAGCTTAG TAGAAGACTT TGGTGAGTAT
1820 1840 1860 1880 1900
ATTATAACTC AATTCTCAT CTCTTTTAGA CTCAAGTGTT TAAGGAACAA ATTGAGCAGC TCAAGCAAAA GGTAAGGTAG TTTTATGAG TGTATATAAA
TAATATTGAG TTAAGAGATA GAGAAAATCT GAGTTCACAA ATTCTTGTGT TAACCTGTCG AGTTCTGTTT CCATTTTCATC AAAAATACTC ACATATATTT
1920 1940 1960 1980 2000
CAGATATAAG TATGTATGCA AATTGTGTAA TATTCCAAGT AAGTAAGCCT CTTGTGCTTG CTTTTTACAA ATTGGAACTC AAAACTTTTG CAGGAGAAAG
GTCTATATTC ATACATACGT TTAACACATT ATAAGGTTCA TTCAATTCGGA GAACACGAAC GAAAAATGTT TAACCTTAGA TTTTGAAAC GTCTCTTTT
2020 2040 2060 2080 2100
CTCTAGCTGC AGAAAACGAG AAGCTCTCTG AAAAGGTATA ATATATTCTT ATGGGTCTCA AGTTAGGGTT GCACATTCTG TTTTATTTC GGTAAAGATA
GAGATCGACG TCTTTTGCTC TTCGAGAGAC TTTTCCATAT TATATAAGAA TACCCAGAGT TCAATCCCAA CGTGTAAAGCA AAAAATAAG CCAATTTCTAT
2120 2140 2160 2180 2200
AGAAAGTTGG GGTTCTTTT GGGGGTTATT AGGTTAGGAG AGTCCTTACT AGTTTTTCTT GGTATCTTC AATCATCAAC CTCTTTTAA TTATGTATTG
TCTTTCAACC CCAAGAAAAA CCCCCAATAA TCCAATCTCT TCAGGAATGA TCAAAAAGAA CCAATAGAAG TTAGTAGTTG GAAGAAATTA AATACATAAC
2220 2240 2260 2280 2300
TTCTATATAT CTCTAATTT GCATCTATTA ATTTTGTGTA ATAACTCTAT TTGAATGCAG TGGGATCTC ATGAAAGCGA AGTTTGGTCA AATAAGAAATC
AAGATATATA GAAGATTAAA CGTAGATAAT TAAAACACAT TATTAAGATA AACTTACGTC ACCCTAGAG TACTTTCTGT TCAAAACAGT TTATCTTATG
2320 2340 2360 2380
AAGAAAGTAC TGGGAAGAGT GATGAAGAGA GTAGCCCAAG TCTTGAAGTA GAGACGCAAT TGTTCAITGG GTTACCTTGT TCTTCAAGAA AG
TTCTTTCATG ACCTTCTCCA CTACTTCTCT CATCGGGTTC AAGACTTCAT CTCTGCGTTA ACAAGTAACC CAATGGAAAC AGAAGTTCTT TC
```

SEQ ID NO:45

Arabidopsis AGL22 genomic sequence

```
-2981 -2961 -2941 -2921 -2901
TACAAGTCAT CGCCGCCGTC GTCAATTTCG GGATCCGGCG AGAAACTGAA CCAAAATAAT ACTTATTTTA CTCGTAAGGA AAATTTGGGC CTAATAAAAG
ATGTTTCAGTA GCGGCGGCAG CAGTAAAAGT CCTAGGCCGC TCTTTGACTT GGTTTTATTA TGAATAAAAT GAGCATTCCT TTTAAACCCG GATTATTTTC
-2881 -2861 -2841 -2821 -2801
CCCAATAATA ATAAAAAGCC CATTAGGGAC TCCGCTTTAT GATAACGGTG ACTGTAGTTT CTTTGTATGT TCAGAGAGAG TGTGTAGTGT AGGGACTGTG
GGGTATTAT TATTTTTCGG GTAATCCCTG AGGCGAAATA CTATTGCCAC TGACATCAAA GGAACACAC AGTCTCTCTC ACACATCACA TCCCTGACAC
```

5	-2781	TAGAAAGAAA ATCTTTCTTT	GAAGCCTAAA CTTCGGATTT	ATGGCTAAAA TACCGATTTT	-2761	GGTTAGGTGC CCAATCCACG	AATGTTTTCAT TTACAAAGTA	-2741	TAGAGAGGCT ATCTCTCCGA	TGGAAGCTGTT ACCTTGACAA	-2721	AAGGGAAAGG TTCCCTTTCC	TCACGAGTCG AGTGCTCAGC	-2701	TCTACTCATATA AGATGAGTAT
10	-2681	AAAACCTCTGA TTTTGAGACT	CACCTTTGACC GTGAAACTGG	AATCAAAACT TTAGTTTGA	-2661	CAAAGACCTC GTTTCTGGAG	ACCAGTTGTG TGGTCAACAC	-2641	TCACGTGCGC AGTGCACGCG	CTCTAAACAC GAGATTTGTG	-2621	TATTTCAATTT ATAAGTTAAA	CAAATATAAA GTTTATATTT	-2601	TGATTCATGCG ACTAAGTAGC
15	-2581	GGTTCCAAAC CCAAGGTTTG	GCCAAATTGAT CGGTTAACTA	GGATGTTCTA CCTACAAGAT	-2561	CCAAATTTAA GGTTTAAAT	TCTACTTTTA AGATGAAAT	-2541	CCAAACCATG GGTTTGGTAC	ACAAATATGA TGTTTATACT	-2521	ATAAACATTA TATTTGTAAT	CTTGATAATA GAACATATTAT	-2501	ATTTTGTGAGC TAAACACTC
20	-2481	TGAACAAACT ACTTGTTTGA	TTTTTTTTTT AAAAAATAAA	CGAAACCAAA GCTTTGGTTT	-2461	CCAAGCTGAA GGTTCGACTT	AAAAACTCAA TTTTTGAGTT	-2441	CGATTTTCTT GCTAAAAGAA	TGTTTTAAAT ACAAATTTTA	-2421	ACGTTAGAAA TGCAATCTTT	GGAATATGTA CCTTATACAT	-2401	TTATGCCGAA AATACGGCTT
25	-2381	ATAAGTAATA TATTCATTAT	TCGATCAGGC AGCTAGTCCG	CACCTCTCTT GTGGAGAGAA	-2361	ATAGTTATTC TATCAATAAG	TCCTAGCAAC AGGATCGTTG	-2341	TTTAACCACT AAATTTGGTA	AGAAGGTTTT TCTTCCAAAA	-2321	GTTTTCTAGT CAAAAGATCA	GTTTCTTAAT CAAAAGATTA	-2301	ATACGTCATC TATGCAAGTAG
30	-2281	AAAATTTTCA TTTTAAAGT	AAAAATACTA TTTTTATGAT	CATTTTGTGT GTAAAAACAA	-2261	TTAAAACTT AATTTTGA	CCATAATTCC GGTATTAAGG	-2241	ATTACTCGTA TAATGAGCAT	GAACACAAAC CTTGTTGTG	-2221	GCAAACCATTA CGTTTGGTAT	TTAATATTTT AATTATAAAA	-2201	GTTGTCAACA CAACAGTTGT
35	-2181	AAAATTTCAA TTTTAAAGTT	ATTATAATTC TAATATTAA	AACTATATTT TTGATATAAA	-2161	GCTTGATTAC CGAACTAATG	CCAATTAGAT GGTTAATCTA	-2141	AGAAAAGAGT TCTTTTCTCA	TAAAGAAGAA ATTCTTCTTT	-2121	AAGAAAAGAG TTCTTTTCTC	TTTACAGTAA AAATGTCAAT	-2101	ATTAACGCAA TAATGCGT
40	-2081	ACCATAATTA TGGTATTAA	TATTTAACAC ATAAATTGTT	CGTATTAAAT GCATAATTAG	-2061	ACATCAACCA TGTAATTGTT	TATGACTTTT ATACTGAAAA	-2041	TTACCGTTTG AATGGCAAAC	CAACTTCATA GTTGAAGTAT	-2021	ATTTCATATAG TAAGTATTAT	TATCATAATA ATAGTATTAT	-2001	AATTCGCAAT TTAAGCGTTA
45	-1981	AATACAACAC TTATGTTGTG	AAGAGTTTCG TTCTCAAAAG	TCGGAAGAGT AGCCTTCTCA	-1961	AAATAATACT TTTATTATGA	CAATAGGGG GTTTATCCCC	-1941	GTGAGTGATA CACTCACTAT	CGAGCCACAT GCTCGGTGTA	-1921	GTATTTCTTGA CATAGAAGAT	AGGGTAGATT TCCCATCTAA	-1901	ATTGCAAACT TAACGTTTGA
50	-1881	TGGAGTAATA ACCTCATTAT	AAGAGAAGAA TTCTCTTCTT	GAATGGGTTT CTTACCCAAA	-1861	GTAGTAGTTG CATCATCAAC	CGTGGAGTAT GCACCTCATA	-1841	CTTTATTGGG GAAATAAACT	GTAAAACTTT CATTTTGAAA	-1821	AATTTAGAAA TTAAATCTTT	TAAAATCTGT ATTTTAAAGC	-1801	TACGGAACAAT ATGCTGTGTTA
55	-1781	GGATCGTGTC CCTAGCACAG	CCAATCAGAT GGTTAGTCTA	TTCTTGTGGC AAGAACCCCG	-1761	TGCTTCGGGT ACGAAGCCCA	CTGGTTTGGG GACCAAAACC	-1741	GTCCCTTTGA CAGGGAACCT	AAAATTTTAG TTTTAAATC	-1721	TGGTGCACAC ACCAGCTGTG	TTTTTATTTT AAAAATAAAA	-1701	ACTCTGGCTC TGAGACCCGAG
60	-1681	GTGCCTCGAG CACGGAGCTC	GGTCCCTCTA CCAGGGAGAT	TTCACTGTTT AAGTGACAAA	-1661	CTTCGTATGA GAAGCATACT	AGGTATGCTT TCCATACGAA	-1641	AAACATTATT TTTGTAAATA	TTATTTTTAA AATAAAATTT	-1621	AAACCCCTTA TTTGGGAAT	ATTTTATTTT TAAAATAAAA	-1601	CTTACCTTTA GAATGGAAAT
65	-1581	ATCACGGTTT TAGTGCCAAA	TGTAAATTGC ACATTTAACG	TTTTTAGTCT AAAAATCAGA	-1561	ATGGAATGAT TACCTTACTA	GATTGTGGCG CTAACACCGC	-1541	ATTGAAATCA TAACCTTTAGT	TATGTTTGGT ATACAAACCA	-1521	TCTGTTGTGT AGACAAACAC	ACGTTGGTGA TGCAACCACT	-1501	AGTATATGTG TCATATACAC
70	-1481	ATTTGTAATG TAAACATTAC	TTGAGCTTAT AACTCGAATA	GTATTAAAT CATAATTTTA	-1461	GTTAAATGAT CAATTTACTA	AAATAACCTC TTTATTGAG	-1441	GTAAGAAAGT CATTTCTTCA	GATTTCAATTT CTAAAGTAAA	-1421	AAATTTTATT TTTAAATAAA	TTGAGTTACA AACTCAATGT	-1401	TATTGCAATTG ATAAGTTAAC
75	-1381	GTTTTATAAA CAAAAATATTT	AAAATCTTTC TTTTATGAAG	AGTGATGATT TCACTACTAA	-1361	GATACCCCA CTATGGGGGT	TTGTGTGTGT AACACACACA	-1341	AATGTTTACT TTAAACAATGA	GGGATTGAAC CCCTAACTTG	-1321	AAAATTTTAT TTTTAAATAA	TGTGCATGAC ACACGTAAGT	-1301	AAACTTTCCA TTTGAAGGT
80	-1281	AATTAGTGCA TTAATCACGT	TAGATTGTAA ATCTAACATT	TTGTATAATG AACATATTAC	-1261	GACTACATGT CTGATGTACA	ATCTGAGTAG TAGACTCATC	-1241	ATATGGTTCA TATACCAAGT	TTAGGTTACA AATCCAATGT	-1221	AACCTCTTTT TTGAGAGAAA	TTTTAAGGACA AAATTCCTGT	-1201	CAATTTTTCG GTTAAAAAGC
85	-1181	ACAAGTTATA TGTTCAATAT	TGCCACATGA ACGGTGACT	TTGACTACTA AACTGATGAT	-1161	AATTTTCAAA TTAAAGTTT	AATTATTGCA TTAATAACGT	-1141	CTAATGTCTT GATTACAGAA	TGAAATTAAC ACTTTAATTG	-1121	AAATTTATTT TTTAATAAAA	GTCATTTCCG CAGTAAAGGC	-1101	AGTTGGATTTC TCAACTTAAG
90	-1081	TTACAAACCA AATGTTTGGT	AGGCCGAAGT TCCGCTTGA	CACAAACTTA GTGTTGAAT	-1061	TTTCTTTTCA AAAGAAAGTC	TAAAAACAAA ATTTTGTGTT	-1041	ACATTGTCTC TGTAACAGGA	CAGAAAAATT GTCTTTTAA	-1021	CTGAAATGTC GACTTTACAG	ATCTTCCCAA TAGAAGGGTT	-1001	ATGTTTTCAT TACAAAAATG
95	-981	ATAAATAAAA TATTTATTTT	ATAATATACA TATTATATGT	GTTGATATTA CAACTATAAT	-961	TTTTGTCTT AAAACRAGAA	TCTGAATTTT AGACTTAAAA	-941	GTTATGAGGT CAATACTCCA	ACCATTACCA TGGTAATGGT	-921	TATAGTACGT ATATCATGCA	AGATTACAAA TCTAAATGTT	-901	AAATGAAAT TTTACTTTTA
100	-881	ACGTTGTAGC TGCAACATCG	CCTTGATGTT GGAACACTAA	CTTCAGGTCT GAAGTCCAGA	-861	TCTAGTTAGT AGATCAATCA	TTTTGCAGTA AAAACGTCT	-841	AATACCAACC TTATGGTTGG	AATTAGTTAC TTAATCAATG	-821	AAGGAGTATA TTCCTCATAT	AGTGAACAAA TCACTTGTTT	-801	GTGAGACAA CACTCTGTTG
	-781	TCATTTTATG AGTAAATAC	CTTCCCTATA GAAGGATAT	AAAAGAAATT TTTTCTTTAA	-761	CCCCTGAC GGGTTGACTG	CCAAACACAC GGTTTGTGTG	-741	ACTTCTCTC TGAAGAGAAG	TCTCTCTCAT AGAGAGAGTA	-721	CTCATTTGGAG GAGTAACCTC	ACTTATAAAT TGAATATTTA	-701	CCTATTACCT GGATAATGGA
	-681	CACCATATCC GTGGTATAGG	AATAACCAACC TTATTGGTGG	ACACACAGAC TGTGTGCTCG	-661	CAATATCCAA GTTATAGGTT	AAAAAAACT TTTTTTTGA	-641	AAAACTAAAA TTTTGATTTT	ATATAATATA TATAATTATAT	-621	TATCGTTTTC ATAGCAAAAG	TTTCCAAAAA AAAGGTTTTT	-601	TAATCATTTA ATTAGTAAAT
	-581	AGAAACCCCA TCTTTGGGGT	TCATCTTGAT AGTAGAAGTA	AGTATTATAA TCATAATATT	-561	AATTAATAAA TTAATTATTT	CCTCTCCCTG GGAGAGGGAC	-541	AAAAATATCT TTTTATAGAG	ATCCTTCACC TAGGAAGTGG	-521	AATCAAAACC TTAGTTTTGG	TTCTCATGTC AAGAGTACAG	-501	TTCTTCTCTC AAGAAGAGAG
	-481	CTCGACCTTT GAGCTGAAAA	GAGGTGAAAA CTCCACCTTT	ATTAAATATA TAATTTATAT	-461	TTCCCTTAGC AAGGGAATCG	TTTTTTTCTC AAAAAAAGAG	-441	CTTTAGTTTT GAAATCAAAA	CTTCTTCTC GAAGAAGAAG	-421	TTGAGTTTTT AACTCAAAAA	TTTCTTTTGA AAAGAAAACT	-401	TCCTCTCTAA AGGAGAGATT
	-381	TTTCTTGTG AAAGGAACAA	GATTCATCGA CTAAGTAGCT	CTAGATCTAA GATCTAGATT	-361	TTCTTCTCAC AAGAAGAGTG	AAAAGACTGA TTTTCTGACT	-341	GTGTGTTCTT CACACAAGAA	TCTTTCAAAAT AGAAAGTTTA	-321	CTTTCAAAAA GAAAGTTTTT	CTAGGGTTTT GATCCCAAAA	-301	TACTGTCTTC ATGACAGAAC
	-281				-261			-241			-221		-201		

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100

AAATCATATT TATTCTTCTA AATTTAGCAA AAAGAACACG ATTTACTTTC CATTTTCAGTC GTCTTGTCCAC TCTCTCTCTC TTCTTTAAAG TCTCCCTTTT
TTTAGTATAA ATAAGAAGAT TTAAATCGTT TTCTTTGTGC TAAATGAAAG GTAAAGTCAG CAGAACACGTG AGAGAGAGAG AAGAAATTTT AGAGGGAAAA

-161 -161 -141 -121 -101
TAGCAAAAAT TCTCTCTCTC ACAAATTTA TTCTCTCTGG CTCTCTCTCC ATCTCTCTCTC TTACTCTCTCT CTTTAATCAT CTCTCACTCT
ATCGTTTTTA AGAGAGAGAG TGTTTTAAAT AAAGGAGACC GAAGAAGAAG GAGGAGGAGG TAGAAGAGAG AAATAGAGAT GAAATTAGTA GAGAGTAGA

-81 -61 -41 -21 -1
TGAATCTTGA TCCATCAAAA TCAATCCCGT TCTCGAAGA TCCATTAAAA TCAAAACCTA AGCTCTCTCT CTTCCTCTCTA GGGTTTTTTT GTTCGTGTGT
ACTTAGAACT AGGTAGTTTT AGTTAGGCGA AGAGCTTTCT AGGTAATTTT AGTTTGTGAT TCGAGAGAGA GAACGAAGAT CCCAAAAAAA CAAGCAACAC

20 40 60 80 100
ATGGCGAGAG AAAAGATTCA GATCAGGAAG ATCGACAACG CAACGCGCAG ACAAGTGACG TTTTCGAAAC GAAGAAGAGG GCTTTTCAAG AAAGCTGAAG
TACCGCTCTC TTTTCTAAGT CTAGTCTCTC TAGCTGTGTG GTTGCCGCTC TGTTCACCTG AAAAGCTTTG CTTCCTCTCC CGAAAAAGTT TTTGCACTTC

120 140 160 180 200
AACTCTCCGT TCTCTGCGAC GCCGATGTGC CTCTCATCAT CTCTCTTCC ACCGGAACAC TGTTCGAGTT CTGTAGCTCC AGGTCTTTCT TTCTCTCTCT
TTGAGAGGCA AGAGACCGCT CGGCTACAGC GAGAGTAGTA GAAGAGAAGG TGGCCTTTTG ACAAGCTCAA GACATCGAGG TCCAGAAAAA AAGAGAGAGA

220 240 260 280 300
AACTCTCCCT TCTATAGATT TCTCATAACT CATCGAAGGA ATCTGTGTCTA GATCCAGACA AAAAACTTTA AAGAGTTTTT AGATGTATAT CTGATACAT
TTGAAGGGAG AGATATCTAA AGAGTATTGA GTAGCTTCTCT TAGAACAGAT CTAGGTCTGT TTTTGAATTT TTCTCAAAAA TCTACATATA GACTATGTAT

320 340 360 380 400
GGAGTTTTACT GTATCAATCT TTATAGGACC ACTAACTATT TATATAATTA AAATAGTTGT TAGAAAAATT AATCATGACC ATAAATGACA TATATAAAGT
CCTCAAAATGA CATAGTTTGA AATATCTCTG TGATTGTATA ATATATTAAT TTTATCAACA ATCTTTGTAA TTAGTACTGG TATTACTCTG ATATATTTCA

420 440 460 480 500
GTATAGTAAA ACTCTGTATT TAGATAAATT AAGGTATCTA ACTACGGTAA TATTCAAAAA GATGTAAATC TGGATATGCA TATATGTATA TTATTAGTAT
CATATCATTT TGAGACATAA ATCTATTTAA TTCCATAGAT TGATGCCATT ATAAAGTTTT CTACATTTAG ACCTATACGT ATATACATAT AATAATCATA

520 540 560 580 600
ATAAATACAT GCTCTATAGT AGGTATTTGT GTCAACCATG TATAAATCTA TGTATATAGA TATGTGTGTA TGATATGTTT AAGCCGTCAA TGTCAATATT
TATTTATGTA CGAGATATCA TCCATAAACA CAGTTGGTAC ATATTTAGAT ACATATATCT ATAACACCAT ACTATACAAA TTCGCGAGTT ACAGTATAAA

620 640 660 680 700
ATATAGAAAT ATGTGGGTAC CATAACATGA GGAAGTATCT ATATGTGTGG ATGTATAAAG CTTCCTCTTT GAAGAAGTAA TCTAAAAATA ATATATATAT
TATATCTTTA TACACCATG GTATTGTACT CCTTCATAGA TATACACACC TACATATTTT GAAAGGGAAA CTTCCTTCATT AGATTTTTAT TATATATATA

720 740 760 780 800
ATATATGTAT ATGTATAGAT ATGTGTGAAT CTTTATTAGT GTTGGGAAAA GTCAITTTAGA GAGATATTAT TGATATTAGG GATCTAAAAA GACTTATCGT
TATATACATA TACATATCTA TACAACCTTA GAAATAATCA CAACCTTTTT CAGTAAATCT CTCTATAATA ACTATAATCC CTAGATTTTA CTGAATAGCA

820 840 860 880 900
ATTACAGAGA TACGATTTTG GATTTTGTAC CCACTAGTTA TCAGCTCAGT TCCTATCTTC GGGGACATAC ACACITTCAC AGATAATTGT GTATATATGT
TAATGTCTCT ATGCTAAAAA CTAAAAAAGT GGTGATCAAT AGTCGATCA AGGATAGAAG CCCCTGTATG TGTGAAAGTG TCTATTAAAC CATATATACA

920 940 960 980 1000
AACTGAAAC GATAGTGTTA ACATGAAATA ATGTACATGT TTGGGATTAA ATGTGTTTTT TGGATTGGT TTGCATCTTT TGATTTTGA TTTTGGTATA
TTGACTTTTG CTATCACAAT TGTACTTTAT TACATGTACA AACCCTAATT TACACAAAAC ACCTAAACCA AACGTAGAAA ACTAAAATCT AAAACCATAT

1020 1040 1060 1080 1100
TTGTGCGTGT TTACATATGC ACATTGTAA TATCAACAGT ATAGTTGTTT ATAATAAGTT ATTTATTGGA ATGTGTTTAT ATTATGAAGC ATGAAGGAAG
AACAGCCACA AATGTATACG TGTAACAATT ATAGTTGTCA TATCAACAAA TATTATTCAA TAAATAACCT TACACAAATA TAATACTTCG TACTTCTCTC

1120 1140 1160 1180 1200
TCCTAGAGAG GCATAACTTG CAGTCAAAGA ACTTGAGAAA GCTTGATCAG CCATCTCTTG AGTTACAGGT TAGCTACATT CTCGAAACGA CCACACATT
AGGATCTCTC CGTATTGAAC GTCACTTTCT TGAACCTCTT GGAACCTAGT GGTAGAGAAC TCAATGTCCA ATCGATGTAA ATCGATGTAA GAGCTTGTCT GGTGTGTAAA

1220 1240 1260 1280 1300
TCTTTCCCGA TTTCTGTAAC TTGCAAAATC GAGTATTACT CCGTTGAATT ACCAATATGT TTTAGATTGT TGTATTTATT GACCAAGAAAT CTCTTAAAC
AGAAAGGGCT AAAGACATTG AACGTTTTAG CTCATAATGA GGCAACTTAA TGGTTATACA AAATCTAACA ACATAAATAA CTGTTCTTA GAGAATTTTG

1320 1340 1360 1380 1400
TTTGTATTAA TAGGTACAAA ACITTTATAT ATTGCATATG ATTAATTAGA CTCGATCCAT GTAGTAGTCA TGTAGAGTAG TCCTGTGTAG AGAGTTGAGC
AAACATAATT ATCCATGTTT TGAATATAA TAACGTATAC TAAATTAATCT GAGCTAGGTA CATCATCAGT ACATCTCATC AGGACACATC TCTCACTCG

1420 1440 1460 1480 1500
TTTAGATCAT TATGGATATG ATTAAGAGCT TAAATCAATG TTTTATTCTG TTAGCTGGTT GAGAACAGTG ATCAGCGCCG AATGAGTAAA GAAATTGCGG
AAATCTAGTA ATACCTATAC TAATCTCGA ATTTAGTTAC AAAATAAGAC AATCGACCAA CTCTTGTCCAC TAGTGGCGGC TTACTCATTT CTTTAACGCC

1520 1540 1560 1580 1600
ACAAGAGCCA CCGACTAAGG TACGTTATAT ATGTATATTC TATGACTTTT GAACTAACTA TCATTTTCTA ACTAATTTTT TTTTGTATCA ACCACTATCA
TGTTCTCGGT GCGTGATTCC ATGCAATATA TACATATAAG ATACTGAAAA CTGTATTGAT AGTAAAGAGT TGATTAAAAA AAAAAGTAGT TGGTGATAGT

1620 1640 1660 1680 1700
TTTTCTAACT GTGTGTTTAC ATGATCATAT ATAGGCAAAAT GAGAGGAGAG GAACTTCAAG GACTTGACAT TGAAGAGCTT CAGCAGCTAG AGAAGGCCCT
AAAAGATTGA CACACAAATG TACTAGTATA TATCCGTTTA CTCTCTCTCT CTGGAAGTTC CTGAACTGTA ACTTCTCGAA GTCGTCGATC TCTTCCGGGA

1720 1740 1760 1780 1800
TGAAACTGGT TTGACGCGTG TGATTGAAAC AAAGGTTGTT AAGAAAAATTA CTGATACCA TGTATAAGTT TCTCTAAGCT TACGAGTATG CAATTTACTA
ACTTTGACCA AACTGCGCAC ACTAACTTTG TTTCACACAA TTCTTTTAAAT GAACTATGGT ACATATTCAA AGAGATTGGA ATGCTCATAC GTTAAATGAT

1820 1840 1860 1880 1900
ATACGAGATG TGTTTGCGAG GTGACAAGAT TATGAGTGAG ATCAGCGAAC TTCAGAAAAA GGTAATAATT AACCAGAAATA ACGTTTATTC TTTACTTGAT
TATGCTCTAC ACAAACGTCT CACTGTTCTA ATACTCACTC TAGTCGCTTG AAGTCTTTTT CCATTATTAA TTGGTTTTAT TGCAAAATAG AAATGAACATA

1920 1940 1960 1980 2000
GATTTCAATA TTAATTTTGG CAGTTTCAAG ATCCAAAAAT TTCACTCTCT TCTCTTTTTT TTTGGTGTTC AGGGAATGCA ATTGATGGAT GAGAACAAGC
CTAAAGTTAT AATTAAAAAC GTCAAAAGTTT TAGGTTTTAA AAGTAGAAGA AGAGAAAAAA AAACCACAAAG TCCCTTACGT TAACACCTTA CTCTTGTTCG

2020 2040 2060 2080 2100
GGTTGAGGCA GCAAGTATGT GTCTTACCTC CTCTGTTGAT AACAATATCC TTTCTTTTGT CTACCATTAA CGTACACACC CCTAAATTTA ATCCCCAGTT
CCAACCTCCG CGTTTATACA CAGAATGGGA GAGACAACTA TTGTTTAGGG AAAGAAAAA GATGGTAATT GCATGTGTGG GGATTTAAAT TAGGGGTCAA

2120
GTCTACAACA CATATGTTTG ATCATACTGT GAGA
CAGATGTTGT GTATACAAAC TAGTATGACA CTCT

SEQ ID NO:46

Arabidopsis AGL24 genomic sequence

-2981 -2961 -2941 -2921 -2901
AGACTTACAA TAACCTCATC AAGCAACTCA TACACGAGCA CAAAGTTTTT CCTGAATGAA TCTTCATTCA GAACACCAAG ATAATCCTTA ATAACACGAG
TCTGAATGTT ATTGAAGTAG TTCGTTGAGT ATGTGCTCGT GTTTCAAAAA GGACTTACTT AGAAGTAAGT CTTGTGGTTC TATTAGGAAT TATTGTGCTC

-2881 -2861 -2841 -2821 -2801
CAATCCTTTG TAGAAGCTCC AAAACAAGAG AGGGTGACAC GTTAACTCTC GTTGTGCGAA CAAAATATAG ACCAACCAACC TTGACATGGA AGTAGITCAC
GTTAGGAAAC ATCTTCGAGG TTTTGTCTC TCCCACTGTG CAATTGAGAG CAACAGCGTT GTTTTATATC TGGTTGTGG AACTGTACCT TCATCAAGTG

-2781 -2761 -2741 -2721 -2701
GCCATCGACA TTCTATAAGC ACAAAAAATA AGTTAGATGA AATCATTACA GCTCACAACC AAACAGAAAG TATAATACCT ACAAGATAG GTGGCGCCTC
CGGTAGCTGT AAGATATTCG TGTTTTTTAT TCAATCTACT TTAGTAATGT CGAGTGTGG TTTGTCTTTC ATATTATGGA TGTTCCTATC CACCGCGGAG

-2681 -2661 -2641 -2621 -2601
TGCATTGCCA TCCTCCTTCC AGAAGTTGAC TTTACGGAAG AATGTCTCTG TACTTCCTTT GGGTACCTCA GCCCGTCTG TAGCAATAAA ACGTATACAA
ACGTAAACGT AGGAGGAAGG TCTTGAAGT AAATGCCTTC TTACAGAGAC ATGAAGGAAA CCCATGGAGT CGGGCCAGAC ATCGTTATTT TGCAATGTGT

-2581 -2561 -2541 -2521 -2501
TCTTGAACT TGTATTGGAT CCAACCAAAT CGTATAATCT CAAAACAAAT AGCTTTCTTC TACTACATTA CATAAGATA CTCTGCCAA ACTAATTGAA
AGAAGTTTGA ACATAACCTA GGTGGTTTGA GCATATTAGA GTTTTGTTTA TCGAAGAAG ATGATGTAAT GTATGTCTAT GAGACGGGT TGATTAAGTT

-2481 -2461 -2441 -2421 -2401
TAGTTTGTCT ATATTGTAC AATCTGATTT GGAATTCAG CTCAACATAA TTTGTCATCG GATAAGAAAT GTTGGTAGAT CAAACAGATC AATGAGCTTA
ATCAAAACGA TATAAACATG TTAGACTAAA CCTTTAAGTC GAGTTGTATT AAACAGTAGC CTATTCTTTA CAACCATCTA GTTTGTCTAG TTACTCGAAT

-2381 -2361 -2341 -2321 -2301
GAGAAGATTT CAATGGAAAA TTCTCATGAA ACAGTGACAT AAGACTCGAC TCTGAAGAGA AAAAGCAAAA CAGGAAGAAG CAGAGAGGAT CAGATCGAGA
CTCTTCTAAA GTTACCTTTT AAGAGTACTT TGCTACTGTA TTCTGAGCTG AGACTTCTCT TTTTCGTTTT GTCTTCTCTC GTCTCTCCTA GTCTAGCTCT

-2281 -2261 -2241 -2221 -2201
AAGAGAGCTT ACAGTCACGG AAGACGATGT TATCTCCTCG CTGAGATAAG ACGAAGAATT GGGAGATCAT CATCGTTCCT TATAGCGGTG GATTCCGACT
TTCTCTCGAA TGTCAGTGCC TTCTGTCTACA ATAGAGGAGC GACTCTATTC TGCTTCTTAA CCCTCTAGTA GTAGCAAGGA ATATCGCCAC CTAAGGCTGA

-2181 -2161 -2141 -2121 -2101
GTTTCACCGC GAGTTTGGTT AAGTCTACTG ATCGCCGATC GGTCTCGTCT TTTTGTGTGT CTGGTGGTGA GGTGGTTCAC GTTTTACCAT TTGCCGTCGT
CAAAGTGCGC CTCAAACCAA TTCAGATGAC TAGCGGCTAG CCAGAGCAGA AAAACACACA GACCACCACT CCACCAAGTG CAAAATGGTA AACGGCAGCA

-2081 -2061 -2041 -2021 -2001
TATCGTGAAG CTTCTTCATG AGACGGAGGG TTCTGTGTTT TTGTGAATTA TGATTCTCTG TTCTTATATG GGCTATTTT TAAGACATCA ATATGGCCCA
ATAGCACTTC GAAGAAGTAC TCTGCCTCCC AAGACACAAA AACACTTAAT ACTAAGAAGC AAGAATATAC CCGGATAAAA ATTCTGTAGT TATACCGGGT

-1981 -1961 -1941 -1921 -1901
AATTTGGAAC TTGTTATGAG TTAAAGGAAA TAAGTAGTAA GTACTATAAA TGATGGTTCG ATCTCGGAGG AGAAAAAAA AAACATTGTT TACGAGGAAG
TTAAAGCTTG AACATACTC AAATCTCTTT ATTCATCAIT CATGATATTT ACTACCAAGC TAGAGCCTCC TCTTTTTTTT TTTGTAACAA ATGCTCCTTC

-1881 -1861 -1841 -1821 -1801
CAAAATGTA GTTGATATAA AGGTGACAAC ACATAATTTA TTTTGGAAAG TCAAAACTTT GAGGATTAAG CTGACAACGA AGGTTAGTGA AGACTTTCGG
GTTTTACACT CAACTATATT TCCCATGTTG TGTATTAAAT AAAAACCTTC AGTTTTGAAA CTCCTAATTC GACTGTGTCT TCCAATCACT TCTGAAAGCC

-1781 -1761 -1741 -1721 -1701
GATCGAGCAA TCGGGAGATA TACATGAGCC TAGAGGGCTG ACAAGATGAC CAAGCATTC AAATGAAAGG CTTAAGATTT TTCTTTTCT AAACCTCAAGT
CTAGCTCGTT AGCCCTCTAT ATGTACTCGG ATCTCCCGAC TGTCTACTG GTTCTGAAGG TTTACTTTCC GAATTCATAA AAGAAAAAGA TTGAGTTCA

-1681 -1661 -1641 -1621 -1601
AAGAAACACA AGATATATGA AAGGGAACA AGGGTCAACA ACAAGTCTAA GCTTTTAAA CGTGTTAGAT GATTCTTCTT GAACACTATT ACAATTACTG
TTCITTTGTG TCTATATACT TTCCCATGTT TCCCATGTTG TGTTCAGATT CGAAAAATTT GCACAATCTA CTAAGAAGAA CTTGTGATAA TGTTAATGAC

-1581 -1561 -1541 -1521 -1501
TTTAGTTTCA CATTATATG ACCTTGGGAG TCTTCTAGCT CGTCCCAAT ATATTTTCAA CATATTACTA TAAGATCCTA AAGACCAATA ACATTGATCT
AAATCAAGT GTAAATATAC TGGAACCTC AGAAGATCGA GCAGGGTTTA TATAAAGTT GTATAATGAT ATTCTAGGAT TTCTGGTTAT TGTAAGTGA

-1481 -1461 -1441 -1421 -1401
ACACCAAAAA CTCTCACTTT CTGATTTTGC ACTCGCTTTT TTCTCTCCCA TAAACAAAA CAAAGGCTTA CAATACTAAA TCTGTCTCAC ATTCTTAGTG
TGTGGTTTTT GAGAGTGAAA GACTAAAACG TGAGCGAAAA AAAGGAGGGT ATTTGTTTTG GTTCCGAAT GTTATGATTT AGACAGAGTG TAAGAATCAC

-1381 -1361 -1341 -1321 -1301
CTTATTGTTT TTAGTCATAA AGAAGTTAAT CTATACAGA TTGAAGTCTT AAAGTCATCT ATATTACTTT TCACATGTAT CATTATGAGA TGGTACGTTT

GAATAAACAA AATCAGTATT TCTTGAATTA GAATATGTCT AACTTCAGAA TTTCAGTAGA TATAATGAAA AGTGTACATA GTAATACTCT ACCATGCAAA

-1281 -1261 -1241 -1221 -1201
CCCACGAATT TTATCAGTTT AGTTTAATTT TCAGTTGTAC TTGGGGAGAA AAAATTTTACA AGATACTTGT CGGCCATGAT ATCACCCCTAG AGTTACCGGA
GGGTGCTTAA AATAGTCAAA TCAAAITAAA AGTCAACATG AAACCCTCTT TTTTAAATGT TCTATGAACA GCCGGTACTA TAGTGGGATC TCAATGGCCT

-1181 -1161 -1141 -1121 -1101
GTCCGGTGAT ATATCATTTT TAATTAGGGT TAAAACTTAA AAGGGTATAA ATGGCTGATC AAACCCAAAA ATAAAAGATA ATGATGACGG TGGGAGACGA
CAGGCCACTA TATAGTAAAG ATTAATCCCA ATTTGAATT TTCCCATATT TACCGACTAG TTTGGGTTTT TATTTCTAT TACTACTGCC ACCCTCTGCT

-1081 -1061 -1041 -1021 -1001
GTGATCTTAT CAGGTGTCGC ATCTAGCATA TATAGTGAA AGACTATAAA AAAGACATGA AATATTTAAT AGACACAAC TTTGTAATA ACCAAAACCA
CACTAGAATA GTCCACAGCG TAGATCGTAT ATATCCACTT TCTGATATTT TTTCTGTACT TTATAAATTA TCTGTGTGA AACATTATT TGGTTTTGGT

-981 -961 -941 -921 -901
AAAAGGTAGA TGAAGTATG AACAGCATCT TCTAATTACG AATAAAAAA GTAACCAAC TTTCTTTCCA TTAGAATTGG TACGTAGTTC CTGTGTATT
TTTTCCATCT ACTTGACTAC TTGTCTAGA AGATTAAATG TTATTTTCTT CATTTGGTTG AAAGAAAGGT AATCTTAACC ATGCATCAAG GAACACATAA

-881 -861 -841 -821 -801
GTGATTTCTT TCATTTTCCA ATTATGTTTT TTTATTTTAT CATGTACAT TTTTGATAGT GGGTAACTTT TGTATCATTT TATTTGACCT AGCCATATAT
CACTAAAGAA AGTAAAGGT TAATACAAAA AAATAAATA GTACAATGTA AAAACTATCA CCCATTGAAA ACATAGTAAA ATAACTGGA TCGGTATATA

-781 -761 -741 -721 -701
AAATCTATTA ACTTATACGG AGTAGTATTT CAGTCAATTT ATTTTATTTT TGTTTTTAGA TGGGAAGTTA TTCAAACTA GACTAAACA GTAAACTAG
TTTAGATAAT TGAATATGCC TCATCATAAA GTGCAGTAAA TAAAATAAAA ACAAATCTT ACCCTTCAAT AAGTTTGTAT CTGATTTTGT CATTTTGATC

-681 -661 -641 -621 -601
GAAACCCGCT ACTGAATAAA GTTACAATTC CACATTATTC CATGACAGAC TAATTGAATT AGAAGGTTAG GTAAATTATT AAATCATAC TGTAGCAGTC
CTTTGGCGA TGACTTATTT CAATGTTAAG GTGTAATAAG GTACTGTCTG ATTAACITAA TCTTCCAATC CATTTAATAA TTTAGTATTG ACATCGTCAG

-581 -561 -541 -521 -501
TCTTCGTCTG GCAGCTCAGT CAGACAAAAC ACAAGTGTG TTTATGTGTT ATTTTAAATG ATTATAGTTT GGGAAAAAGA CATAATCAAA AGGGATACAA
AGAAGCAGAC CGTCGAGTCA GTCTGTTTTG TGTTTCACAC AAATACACAA TAAAATTTAC TAATATCAAA CCTTTTCTCT GTATTAGTTT TCCCTATGTT

-481 -461 -441 -421 -401
AACATATGGC CCATTGATAA GTATAGATCA CTGTTTAGCT AAAAAAGCA GACTCTTTT TCCAATCTTG AACACAAACA CAGTCACCAT CTCTCTCTCT
TTGTATACCG GGTAACTATT CATATCTAGT GACAAATCGA TTTTTCGT CTGAGAAAAA AGGTTAGAAC TTGTGTTTGT GTCTAGTGA GAGAGAGAGA

-381 -361 -341 -321 -301
CTTTCTCTCT CACTCACACA TTAGGGAGTA AACAGTACC AGAAAAACCT TTTTATCTT CTCACAAATT TAATAAAGTG GGTGCTGAGA TTGAATAACG
GAAAGAGAGA GTGAGTGTG AATCCCTCAT TTGTCTAGG TCTTTTGA AAAAATAGAA GAGTGTTTAA ATTATTTTAC CCACGACTCT AACTTATGTC

-281 -261 -241 -221 -201
TAATCCAAGA TCCTCCAAC CACAGAAAG TAAAGCTGT GAATCTGTG TCTTTCTTCT TAAGCAAAGT GTTTGATGAA TTCATCTAGT CCTGTCCATT
ATTAGGTTCT AGGAGGTTGA GTGCTTTTCC ATTTCTGACA CTTAGACACA AGAAAGAAGA ATTCGTTTCA CAAACTACTT AAGTAGATCA GGACAGGTAA

-181 -161 -141 -121 -101
CTTTTGCTTC TCATGGTTTA TGGATCTGAT CTCTCTTCT CTCTCTCTCT AGCCATTAGG GTTTCCTAAG AATATTATAT AAATCTCTCT TAGCTAACAC
GAAAACGAAG AGTACCAAT ACCTAGACTA GAGAGAAAGA GAGAGAGAGA TCGGTAATCC CAAAGGATTC TTATAATATA TTTGAGAGAA ATCGATTGTG

-81 -61 -41 -21 -1
CGTTCCAATT GGTTCCTTTC TTTGTTCTTG GTCTAAAATC TAAATGGTGT TATGGGTATA GGCAGATTCA AGAACAGTAG TGAAGGAGAG ATCTGGTAAA
GCAAGGTTAA CCAAAGAAAG AAACAAGAAC CAGATTTTAG ATTTACCACA ATACCCTAT CCGTCTAAGT TCTGTCTATC ACTTCTCTCT TAGACCATTT

20 40 60 80 100
ATGGCGAGAG AGAAGATAAG GATAAAGAAG ATTGATAACA TAACAGCGAG ACAAGTTACT TTCTCAAAGA GAAGAAGAGG AATCTTCAAG AAAGCCGATG
TACCGCTCTC TCTTCTATTC CTATTTCTTC TAACTATTGT ATTGTGCTC TGTTCATGA AAGAGTTTCT CTCTCTCTCC TTAGAAGTTC TTTCGGCTAC

120 140 160 180 200
AACTTTTCACT TCTTTGGGAT GCTGATGTTG CTCTCATCAT CTCTCTGCCC ACCGGAAAGC TCTTCGAGTT CTCCAGCTCA AGGTATATTC TATCTTTTGT
TTGAAAGTCA AGAAACGCTA CGACTACAAC GAGAGTAGTA GAAGAGACGG TGGCCTTTTC AGAAGCTCAA GAGTCTGAGT TCCATATAAG ATAGAAAAAC

220 240 260 280 300
TTAGTAGTTG TCTTATTTTT TTCAATCCAT GTTTGTGTTT TTGAGAATAT GGTGGGATAA ATATATTAAG ATATGTATTT AAATGAGATT TTTATTTTCT
AATCATCAAC AGAATAAAAA AAGTTAGGTA CAAACACAAA AACTCTTATA CCAACCTATT TATATAATTC TATACATAAA TTTACTCTAA AAATAAAGA

320 340 360 380 400
CGTTTACTCT CTAAAGTTAA TTATCAGTAG GCTCGGAGAT CTCATGTACG GCATAATTG ATGACCTAAA TTATTATACT TTAAAGTATA GGATTGATGT

GCAAATGAGA GATTTC AATAGTCATC CGAGCCTCTA GAGTACATGC CGTATTAAC TACTGGATT TATAATATGA AATTCATAT CCTAATACA
 420 440 460 480 500
 TTTATTACTT TTATGTATAA CACATCATGT ATTTAATTC GTTTAACATA ATATGGGTTT TTAACGTGTA ATTTTTCAT CATTTCATT TAGACTCATG
 AAATAATGAA AATACATATT GTGTAGTACA TAAATTAAGG CAAATGTAT TATACCCAAA AATTGCACAT TAAAAAGTTA GTAAAAAGTAA ATCTGAGTAC
 520 540 560 580 600
 GTTAAGATTT CTGTACTGGG AAATAAGAGA GCAGAATATT ATAGTGTGAT TTTTGTAAAT TAGGAAAGCA TATGTATATA TGGATACATA GTACTTACCA
 CAATTCTAAA GACATGACCC TTTATTCTCT CGTCTTATAA TATCACACTA AAAACAATTA ATCCTTTCGT ATACATATAT ACCTATGTAT CATGAATGGT
 620 640 660 680 700
 CAATTAGAAT GAATTTCTTT TCCCTTTTTT CATTTGACTT TGTGTAITAC AAAAGTCTTT GACACTGTCA CTGCGTATGA TTGGGGATTA ATCTTTAACC
 GTTAATCTTA CTTAAAGAAA AGGGAAAAAA GTAAACTGAA ACACATAATG TTTTCAGAAA CTGTGACAGT GAACCATACT AACCCCTAAT TAAGAATTGG
 720 740 760 780 800
 ACTCGTTTAG TTTATCTTGG GAAGCATTAC CATAATTGGG AAACGAGTCA TCTGTCTGTA TCGTGATGGC TACTTCTGAT TACTTTTCTT TTATTATAAC
 TGAGCAATC AAATAGAACC CTTCTTAATG GTATTAAACC TTTGCTCAGT AGACAGACAT AGCACTACCG ATGAAGACTA ATGAAAAGAA AATAATATTG
 820 840 860 880 900
 CAAAAAGGCT TCTAATGTAC TTAATTAAAT TTACAAATGT AATATGGACG AAGGAAATGT TTATAAGAAA GATGGATTGT TTGTTGAAAC GTGTAGAATG
 GTTTTCCGA AGATTACATG AATTAATTA AATGTTTACA TTATACCTGC TTCTTTTACA AATATTCTTT CTACCTAACA AACAACTTTG CACATCTTAC
 920 940 960 980 1000
 AGAGACATAT TGGGAAGGTA TAGTCTTCAT GCAAGTAACA TCAACAAATT GATGGATCCA CCTTCTACTC ATCTCCGGGT ATTTTCGATA TCACTTACTC
 TCTCTGTATA ACCCTTCCAT ATCAGAAGTA CGTTCATTGT AGTTGTTTAA CTACCTAGGT GGAAGATGAG TAGAGGCCCA TAAAGCTAT AGTGAATGAG
 1020 1040 1060 1080 1100
 TTTTMTTTT TTGTGGATT TAAACTCTCT GCTCTTTTTA CCAACCCTT CTCTTTTAT CAAACCCTTC TCTCTATAAT ATTATCCGAT GTTCACITTG
 AAAAAA AAAACCTAAA ATTTGAGAGA CGAGAAAAAT GGTTTGGGAA GAGAAAAATA GTTTGGGAAG AGAGATATTA TAATAGGCTA CAAGTGA AAC
 1120 1140 1160 1180 1200
 TTACACGTGT TTGTATAAT TTTTAGCTGT AAGTCTAAAT ATAGAAACAT TGAGTGGCAT ATAATCATT AATCTGAAGC ATCTAATTAA TTGGTTTTAC
 AATGTGCACA AACAAATTA AAAATCGACA TTCAGATTTA TATCTTTGTA ACTCACCGTA TATTAGTAAT TAGAACTTCG TAGATTAATT AACCAAAATG
 1220 1240 1260 1280 1300
 ATATTAATAG CAGAACTCTG AAACGTGTGA CTTTGCATCT AGCAGCTTGA GAATGTAAAC CTCTCCAGAC TAAGTAAGGA AGTCGAAGAC AAAACCAAGC
 TATAATTATC GTCTTAGGAC TTTGACAAC TAAACGTAGA TCGTCGAAC CTTAACATTG GAGAGGTCTG ATTCATTCTC TCAGCTCTCG TTTTGGTTCG
 1320 1340 1360 1380 1400
 AGCTACGGTA TGGCTCCATT GATATGTTAT GCAGATAAAC CTATTTTCAT ATAGGCTATA GCTGTAAGAG ATCATCTATT TCATGTGTGT GGTTTTTTTT
 TCGATGCCAT ACCGAGGTAA CTATACAATA CGTCTATTG GATAAAAGTA TATCCGATAT CGACATTCTC TAGTAGATAA AGTACACACA CCAAAAAA
 1420 1440 1460 1480 1500
 TTTATGTTTT TTCAATGATG TGTGCATGCT ATTTTATAGT TTTAGAATCT ATTTCAATGA AATTGAAGAT ATTTCAATTC ACGTGTAAGT TCGTCAAGTT
 AAATACAAA AAGTTACTAC ACACGTACGA TAAAAATCCA AAATCTTAGA TAAAGTACCT TTAACCTCTA TAAAGTAAAG TGCACATTCA AGCAGTTCAA
 1520 1540 1560 1580 1600
 GTGGCGTGTG TCTTGGAAAT TGATGTTTTT TTTGTAGATT TTAAGAGCTA CTCTAAAAAT TTACAAGAGT TTTGTAATTT TCAATTATGG CCCATTATTC
 CACCGCACAC AGAACCTTTA ACTACAAAAC AAACATCTAA AATTCCTGAT GAAGATTTTA AATGTTCTCA AAACATTAAA AGTTAATACC GGGTAATAAG
 1620 1640 1660 1680 1700
 TCATTAATTC ATTAATAAAA TTATATACAT TACTATCTAT ATCTAGCATA GGTAGTTTTT TTTTCTTTT TCTTTGGTAG ACCTACTGAA CAAATATCTG
 AGTAATTAAG TAATTTTTTT AATATATGTA ATGATAGATA TAGATCGTAT CCATCAAAAA AAAAAGAAAA AGAAACCATC TGGATGACTT GTTTATAGAC
 1720 1740 1760 1780 1800
 ATATATCACT GACTGGATAA ATATCTATAG AGATATTTTT GATAGAAATG AGTGTAAATT TAACGTAAAA CAGGAAACTG AGAGGAGAGG ATCTTGATGG
 TATATAGTGA CTGACCTATT TATAGATATC TCTATAAAAA CTATCTTTAC TCACAATTA AATGCAITTT GTCTTTTGAC TCTCCTCTCC TAGAACTACC
 1820 1840 1860 1880 1900
 ATTGAACCTA GAAGAGTTGC AGCGCTGGA GAACTACTT GAATCCGGAC TTAGCCGTGT GTCTGAAAAG AAGGTTTACT ACTATACATA AACTAATAGC
 TAACTTGAAT CTCTCAACG TCGCCGACCT CTTTGATGAA CTTAGCCGTG AATCGGCACA CAGACTTTTC TTCCAATGA TGATATGTAT TTGATTATCG
 1920 1940 1960 1980 2000
 ATGCATATTT TCCTTAACGT GGCATATAAA TAATAAGCTG TACATATATA AAAGTTTGAC TTTGTTGTTG TTATTGGTAA ATAGGGCGAG TGTGTGATGA
 TACGTATAAA AGGAATTGCA CCGTATATTT ATTATTCGAC ATGTATATAT TTTCAAACTG AAACAACAAC AATAACCATT TATCCCGCTC ACACACTACT
 2020 2040 2060 2080 2100
 GCCAAATTTT CTCACCTGAG AAACGGGTTA GTAGTTAGTA CATACAATTC GTATAACTAA TGGATCATAA GCCTATCTAT AGCTAGTGAC TTTCTTAATA

CGGTTTAAAA GAGTGAACCTC TTGCCCCAAT CATCAATCAT GTATGTTAAG CATATTGATT ACCTAGTATT CGGATAGATA TCGATCACTG AAAGAATTAT

2120 2140 2160 2180 2200

AGTGAACACAG GGATCGGAAT TGGTGGATGA GAATAAGAGA CTGAGGGATA AAGTACGGCT CTAAACCCCTT ATAGATATCA TGGAAATAACC TTAATCTATT

5 TCACITTTGTC CCTAGCCTTA ACCACCTACT CTTATTCTCT GACTCCCTAT TTCTAGCCGA GATTTGGGAA TATCTATAGT ACCTTATTGG AATTAGATAA

2220 2240 2260 2280 2300

TTTTTATGTA TAAGAAAATA TGATGAGGGA ACGTATATTA TATATCGGCA GCTAGAGACG TTGAAAAGGG CAAAACCTGAC GACGCTTAAA GAGGCTTTGG

10 AAAAAATACAT ATTCTTTTAT ACTACTCCCT TGCATATAAT ATATAGCCGT CGATCTCTGC AACCTTTCCC GTTTTGACTG CTGGCAATTT C7CCGAAACC

2320 2340 2360 2380 2400

AGACAGAGTC GGTGACCACA AATGTGTCAA GCTACGACAG TGGAACTCCC CTTGAGGATG ACTCCGACAC TTCCCTGAAG CTTGGGTATA ATTTGTTTAA

15 TCTGTCTCAG CCACTGGTGT TTACACAGTT CGATGCTGTC ACCTTGAGGG GAACTCCTAC TGAGGCTGTG AAGGGACTTC GAACCCATAT TAAACAAATT

2420 2440 2460 2480

CTGAACATAT TTCAAACTTT TTGTTGACAT TTTGTATGTG GATGTTTACT AACTGTTTGT TGGTTAGGCT TCCTCTTGG GAA

GACTTGTATA AAGTTTGAAA AACAACCTGA AAACATACAC CTACAAATGA TTGACAAACA ACCAATCCGA AGGTAGAACC CTT

SEQ ID NO:47

Arabidopsis AGL27 genomic sequence

-2961 -2941 -2921 -2901

CAACCAGCAG CACCAGCTGC AATCAAATCC TTTACGGTTC TTTGAATGTT TAGCGCATTT CCTCCACCGG TATCTTGAAA AGATCAAAAG AAACCTATGA

25 GTTGGTGCCTC GTGCTCGACG TTAGTTTAGG AAATGCCAAG AAACCTTACAA ATCGCGTAAA GGAGGTGGCC ATAGAACTTT TCTAGTTTTC TTTGGTAACT

-2881 -2861 -2841 -2821 -2801

AGAGAAGTAT AACCAAGCAA ATCCACTATT TTCAAAAAGC TATGAAGAGA ACTATAAGCA AGCAAGCGAC TCTAACCAAG AAAGATTGAT ACTTTCAATC

30 TCTCTTGATA TTGTTTCGTT TAGGTGATAA AAGTTTTCG ATACTTCTCT TGATATTCTG TCGTTCGCTG AGATTGGTTC TTTCTAATA TGAAGTTAG

-2781 -2761 -2741 -2721 -2701

TTTGGTAAAG AATCAACGAC TCAATGTTTT TAAATGTTTT TTTTCTCTTT TTGGTTTATG TTAAGCTTCT TGCATTCTTT AATGATGTCT TTATTATACT

35 AAACCATTTT TTAGTTGCTG AGTTACAAAA ATTTACAAAA AAAAGGAAAA AACCATAATC AATTGGAAGA ACGTAAGAAA TTACTACAGA AATAATATGA

-2681 -2661 -2641 -2621 -2601

ATCAAAATTT TGCAACTTTA CCAGCATCTG CAATGATGGG TATATTAGGA GCTGACGCAC ACACCGACCT TGGCGTCGCA GCCATCTCCG GTGGTCTAAA

40 TAGTTTTTAA ACGTTGAAAT GGTCTGAGAC GTTACTACCC ATATAATCTT CGACTGCGTG TGTGGTGGGA ACGGCAGCGT CCGTAGAGGC CACCAGATTT

-2581 -2561 -2541 -2521 -2501

ACGACGAAAG AACCAAAATA AACGAAAGC ATACAAACAA AAAATTACTA AAGAAAGAAA AAAAAAAGG TGGCGCACGT TAGCAAAACG AAATCGGGTT

45 TGCTGCTTTC TTGTGTTTTT TTTGCTTTCG TATGTTTGTG TTTTAAATGAT TTCTTTCTTT TTTTFTTTC ACOCGTGCGA ATCGTTTGGC TTTAGCCCAA

-2481 -2461 -2441 -2421 -2401

TTCCAGGAG AGAAGCGGAT AAGCGCTAAC CGGATATAAA ACCAGCGGAG AATCGGTTT GCTGCACAA ATGCCGCGAT AAGGCATCGT AGCATCCAGG

50 AAGGGTCTCT TCTTCGCCTA TTCCGCATTG GCCTATATTT TGGTCGCCTC TTAGGCCAAA CGACGTGTTA TCGGCGCTTA TTCCGTAGCA TCGTAGGTCC

-2381 -2361 -2341 -2321 -2301

CATAAGCACA ATGCTTGTGT CTTCAATCAG GCGATGAAAA CGTGTTTGGA TTCTCGCTGT CGGATTCAAC AATCTCGCGC CGCTGGGGTT CCGTCGGAAT

55 GTATTCTGTG TACGGAACAA GAAGTAGTCT CGCTACTTTT GCACAAACCT AAGAGCGACA GCCTAAGTGG TTAGAGCGGC GCGCACCCAA GGCAGCTTTA

-2281 -2261 -2241 -2221 -2201

GTTGGTGAAG CTGTAAGGTT TAAGCTGCTA CAACAGAGTG AAGTTGTTTT GACAGCCATT AACATCGACA TTCTTCGAAG CCTCGAACAA GTTTTTTCTT

60 CAACCACTTC GACATTCCAA ATTCGACGAT GTTGTCTCAC TTCAACAAAA CTGTGCGTAA TTGTAGCTGT AAGAAGCTTC GGAGCTTGTT CAAAAAGAA

-2181 -2161 -2141 -2121 -2101

CTCTCTAATC GAGTTAGACT CTGACCCACA CGCTTGGGAT TTTAATAGAG AGCACGTGGT TATTATATCT CGGTCTTATC TTATGGTAA CAGTATCTCAA

65 GAGAGATTAG CTCAATCTGA GACTGGGTGT GCGAACCCTA AAATTATCTC TCGTGCACCA ATAATATAGA GCCAGAATAG AATACCATTT TCATAGAGTT

-2081 -2061 -2041 -2021 -2001

AGACTCAAA CACAAGGTAT TGTGAAATG TTAGAGGCAA TCTAACAATA AATGTATAAT TTGGTTAGCT TAAGCTCATC ATAGAAATGG GCCTTTTATG

70 TCTGAGTTTG GTGTTCCTA ACACCTTTAC AATCTCCGTT AGATTGTTAT TTACATATTA AACCAATCGA ATTCGAGTAG TATCTTTACC CGGAAATACA

-1981 -1961 -1941 -1921 -1901

CACCAAACTT ATTTCAACAC ATAACACAAG AGCCCAACAA ACAACGACTC CTTTCTCCAC CAGAACAAGC ACGACAAAGG CAAGAGAGTT GCAAAAGACC

75 GTGGTTTGGT TAAAGTGTG TATTGTGTTT TCGGGTGTGT TGTTGCTGAG GAAAGAGGTG GTCTTGTTTCG TGCTGTTTCC GTTCTCTCAA CGTTTCTCTG

-1881 -1861 -1841 -1821 -1801

TATAAGATGA TAACAATCGA AAAGATGTAA ATTTTGAGAA AAATCAAAAT AAACAAGAAA GATTTCATTG TTTTTCATT TTTCTCCATT TCTATTCTGA

80 ATATTCTACT ATTGTTAGCT TTTCTACATT TAAAACTCTT TTTAGTTTAA TTTGTTCTTT CTTAAAGTAA AAAAAAGTAA AAAGAGGTAA AGATGAAACT

-1781 -1761 -1741 -1721 -1701

TTTTACATAC TCTATGGGCC AACCAATTTT CAACCTAATG CTTGATAAAA AATGATTCCG TTTTACTATC TCAACAAATT GGGCCTACAA CATCCAATTT

85 AAAATGTATG AGATACCCGG TTGGTTAAAG GTTGGATTAC GAACATTTT TTAATAAGCC AAAATGATAG AGTTGTTTAA AGTTGTTTAA CCCGATGTT GTAGGTTAAA

-1681 -1661 -1641 -1621 -1601

CATGTAGTGA CTTGTTTTTG CCTTTTTCAC ATCTCAACAA ATTGGGTCGT TTGTATTTAA GAAATTGTGA CAGCTTTTAA GACTGAATTT TACTTTATGG

90 GTACATCACT GAACAAAAAC GGAAAAAGTG TAGAGTTGTT TAACCCAGCA AACATAAATT CTTTAACAAT GTGCAAAATG CTGACTTAAA ATGAAATACC

-1581 -1561 -1541 -1521 -1501

CTTTATGCTC TCTTTTTCGG TTTTGATTAA GGGTGAATAT GTAACTGTTT GATACCATCT GATTTTCTTT ATTTTCTTAT TTTCTGTGTT GCAACTATAC

95 GAAATACGAG AGAAAAAGGC AAAACTAATT CCCACTTATA CATTTGACAA CTATGGTAGA CTAACAAAAA TAAAAAATAA AAAGAACACA CGTGTATATG

-1481 -1461 -1441 -1421 -1401

CATCTGAATT CAATTGACAT TTTAGCCAAA TAAAAAGAT TGGTCCACTT GGATGGCTGT AAAAAAGTTT AGTGGAGATA TTTATAGGGC TTGTTGGCAA

100 GTAGACTTAA GTTAACGTGA AAATCGGTTT ATTTTCTCTA ACCAGGTGAA CCTACCGACA TTTTCTTCAA TCACCTTCAT AAATATCCCG AACACCCGTT

-1381 -1361 -1341 -1321 -1301

TCITCACCAA CGGCTATAAT GTTGATCTTT TTAATAATTAA ACTTACCGTT CGACTGTCTT CTCAACGATT TGACAAATTG CCGTTAGATT AGTATTACTG

105 AGAAGTGTTT GCCGATATTA CAACCTAGAA AATTTTAAAT TGAATGGCAA GCTGACAGAA GAGTTGCTAA ACTGTTAATC GGCAATCTAA TCATAATGAC

-1281 -1261 -1241 -1221 -1201

ATTTATTATT AACAAACCCA TTTCTTTTCT TATTTTGAAT TAAGCTAAAT CAGGCCAATA AAAGGGACAA GTAGAGATGG GCTATTCTCT TTTTCTCTCT

110 TAAATAATTA TTGTTTGGGT AAAGAAAAAG ATAAAAAATT ATTGATTTTA GTCCGGTTAT TTTCCCTGTT CATCTCTACC CGATAAGAAA AAAAAAGAAA

5	TTTTTTTTTC AAAAAAAAAG	-1181 TTATGTAGTA AATACATCAT	GAGAAAAGCC CTCTTTTCGG	-1161 TTTATTCCTTA AAATAAGAAAT	GAGCTATCAT CTCGATAGTA	-1141 TTACCAACCCA AATGGTGGGT	TTAACCAGAA AATGGTCTCT	-1121 GCTGAGAAAT CGACTCTTTA	GAAGCAAGCC CTTCGTTCCG	-1101 GAAACGAATT CTTTGCTTAA
10	TGTAGTTTTG ACATCAAAAC	-1081 GACGGTGAAA CTGCCACTTT	TTATATCGGG AATATAGCCC	-1061 CCTTTAATGG GGAAATTACC	GCATGTGAAT CGTACACTTA	-1041 AGAGTTGAGA TCTCAACTCT	GTCTTTTTGC CAGAAAAACG	-1021 CCCAAATAAT GGGTTTATTA	CGTTTAAGGG GCAAATTCCT	-1001 AGTATTGGCT TCATAACCGA
15	CGTTGGTTTA GCAACCAAA	-981 ATATTGGGCC TATAACCCGG	GAAACGAGAT CTTTGCTCTA	-961 TGGGAAGAAG ACCCCTCTTC	AACAATGTCT TTGTTACAGC	-941 GTTTAATCCG CAAATTAGGC	GTTAGGGTCG CAATCCAGC	-921 TGGGCTGATT ACCCGACTAA	CTGGTTCACC GACCAAGTGG	-901 TTTATAGCGT AAATATCGCA
20	AAGCGAACAA TTCGCTTGTT	-881 ACATTGAATA TGTAACCTTT	TGGGAAGGCC ACCCCTCTCG	-861 AAATTAGTTA TTTAATCAAT	CCATCCCTAA GGTAGGGATT	-841 CTCAGTTTTG GAGTCAAAAC	AGACGTAGTA TCTGCATCAT	-821 TGAATGAGCC ACTTACTCGG	ACGGCAGAAC TGCCGCTCTG	-801 CTACGACCTA GATGCTGGAT
25	ACTCGATAAA TGAGCTATTT	-781 GTAATGGTTA CATTACCAAT	CTCTTGGAGA GAGAACTCTC	-761 CGGAAGAAAG GCCTTCTTTC	CACAAAGATT GTGTTCTTAA	-741 TTGATAAGGC AACTATTCCT	TTTCTAGTTG AAAGATCAAC	-721 GTGAAATGGT CACTTTACCA	CGGAAATCGT GTTTTCAGCG	-701 CGGAGAGCCA GCCTCTCGGT
30	TCATAGGAGC AGTATCCCTG	-681 GGGGAGGTGC CCCCTCCACG	TATCTGAATA ATAGACTTAT	-661 TCCCAATGCA AGGGTTACGT	TCAAGACAAG AGTTCTGTTC	-641 ATGGATTTCG TACCTAAGTC	AAAACAAGA TTTGTGTTCT	-621 AATTAAACAA TTAATTTGTT	ACATTTTAAA TGTAATAATT	-601 ATATGCTCTT TATACGAGAA
35	AGTTTATAGT TCAAAATCTA	-581 AATATAATGT TTATATTACA	TTTCAATACC AAAATTATGG	-561 AATTATCTTA TTAATAGAAAT	CACCTGATAGT GTGACTATCA	-541 GGTCAAGTTA CCAGTTCAAT	CTAATCACTT GATTAGTGAA	-521 TTAATAAATT AATTATTTAA	GGTGATAGTC CCACTATCAG	-501 AAACGTATTG TTTGCTAATC
40	AAAATTATCG TTTTAATAGC	-481 ATTTAAAAAT TAAATTTTTA	ATTTGAATTC TAAACTTAAAG	-461 AAAACCATTT TTTTGGTAAA	TAGTGAAAGT ATCACTTTTCA	-441 TTGCATTGTA AACGTAAACAT	GTTTTGATTA CAAACTAAT	-421 TCCGATCAAT AGGCTAGTTA	CTTTAATATA GAAATTATAT	-401 ATTACGTCAA TAAATGACGTT
45	TAATAACTGA ATTATTGACT	-381 AATCCCTGAA TTAGGAACCT	TTAACCCTTA AATTGGCAAT	-361 CCCCTATCAT GGGCTAAGTA	AAGCACTACT TTCGTGATGA	-341 TTCCGATCAA AAGGCTAGTT	AACCAATGAG TTGGTTACTC	-321 ATAAAATAAC TATTTTATTG	TTTTAAACCC AAAAATTTGG	-301 TCCAAATPAA AGGTTTATTT
50	AAGAGAAAAA TTCTCTTTTG	-281 CTTAAAAAAC GAATTTTGTG	AATTTCTGTT TTAAGAGCAA	-261 CGGTGGGGAT GCCACCCCTA	GATGATCGGA CTACTAGCCT	-241 CTCGGACCGG GAGCCTGGCC	TCTAACCGAC AGATTGGCTG	-221 TGGAATPAAA ACCTAATTTT	AGTCTTTAAC TCAGAAATTG	-201 AACGACRAGC TTGCTGTTTC
55	TTAAAAATTT AATTTTTTAA	-181 GCCTCTTAGT CGGAGAATCA	GGCTTCAAAA CGGAAGTTTT	-161 CGCAATCGTT CGGTTAGCAA	TGCTTAAATA AGCGAATTAT	-141 CTATTATTTT GATAATAAAA	CTCTATCTCG GAGATAGAGC	-121 TTTAAACCAA AAATTTGGTT	AAAAAAAACG TTTTTTTTTC	-101 AGTTGGAGGA TCAACCTCCT
60	AAAAAAAAC TTTTTTTTTG	-81 CAAGAAAAAA GTCTTTTTTT	GAATAAAAAAG CTTATTTTTT	-61 CAAAAAGCAT GTTTTTCGTA	TGAGCGTCTC ACTCGCAGAG	-41 CGGAGATTAG GCCTCTAATC	GATTAAATTA CTAATTTAAT	-21 GGGCATPACC CCCGTATTGG	CTTATCGGAG GAATAGCCTC	-1 ATTTGAAGCG TAAACTTCGG
65	ATGGGAAGAA TACCCTTCTT	20 GAAAAATCGA CTTTTTAGCT	40 GATCAAGCGA CTAGTTCGCT	60 ATCGAGAAC TAGCTCTTGT	80 AAAGCAGTCG TTTCGTGAGC	100 ACAAGTCACT TGTTCACTGA	120 TTCTCCAAAC AAGAGGTTTG	140 GACGCAATGG CTGCGTTACC	160 TCTCATCGAC AGAGTAGCTG	180 AAAGCTCGAC TTTCGAGCTG
70	AACCTTCGAT TTGAAAGCTA	120 TCTCTGTGAA AGAGACACTT	140 TCCTCCGTCG AGGAGGCAGC	160 CTGTGTGCGT GACAACAGCA	180 CGTATCTGCC GCATAGACGG	200 TCCGAAAAAC AGGCCTTTTG	220 TCTATGACTC AGATACTGAG	240 TTCCTCCGGT AAGGAGGCCA	260 GACGAGTAAG CTGCTCATTC	280 AAGATACTTT TTCTATGAAA
75	CCTTTTCTGG GGAAAAGACC	220 GTCTCACTCG CAGAGTGAGC	240 ATTTTTGTGC TAAAAACACG	260 TTTTTTACTT AAAAAATGAA	280 TGTTTAAATTA ACAAATTAAT	300 CTTCTCCCAT GAAAGAGGTA	320 ATAGAAGCCT TATCTTCGAA	340 CAAATCTAGG GTTTAGATCC	360 GCTTTTTGAT CGAAAACTA	380 TCCATCAAAAT AGGTAGTTTA
80	CAACTGAGAT GTTGACTCTA	320 TTTCTCCTTG AAAAGAGAAC	340 TTTTCTGTAT AAAAGACATA	360 GAAGATAGCA CTTCTATCGT	380 GATGCGTAAG CTACGCAATC	400 CTTTAACCTA GAAATTGGAT	420 ATTTAAGACT TAAATTTCTG	440 AAACATTTTG TTTGTAAAC	460 ATCGCCAAGA TAGCGGTTCT	480 TATGTTCTTG ATACAGAAC
85	ATGTTCTGTT TACAAGCAAA	420 CGTGTPTTTT GCACAAAAAA	440 TTTTCGTGTT AAAAGCAGAA	460 TTTTTTTTTT AAAAAATAAT	480 TCATTTTAAA AGTAAAAATT	500 ATCATTTTAA TAGTAAAAAT	520 TCTCTTTTTT AGAGAAAAAA	540 TACCTTCATT ATGGAAGTAA	560 TGTGACGAAA ACACTGCTTT	580 TTTAATATTG AAATTATAAC
90	CATGTTAATC GTACAATAAG	520 AAGAAACTTT TTCTTTGAAA	540 TCTACACSTG AGATGTGCAC	560 GTGATTCGTT CACTAAGCAA	580 CTTGATGTGG GAACTACAAC	600 TTTAAGTAAT AAATTCATTA	620 CTTTGTATTG GAAACATAAC	640 CTAGTTCCAT GATCAAGGTA	660 CTGTTGTCCA GACAACAAGT	680 CTTTGAAGCT GAAACTTCGA
95	TCGTTTTTTC AGCAAAAAAG	620 ATATAAGAAA TATATCTTTT	640 CAATATGTTT GTTATACAAA	660 AGATTGTGTA TCTAACCAAGT	680 AATTTTGAGA TTAAACTCTC	700 TTTGTTAATT AAACCATTA	720 ATATTCAATA TATAAGTTAT	740 TTGCAATGCA AACGTTACGT	760 CTTCAAGTAG GAAGTTCATC	780 TTTTGTTGAG AAAACAATC
100	AGATTATTTG TCTAATAAAC	720 GGGTAGTGG CCCAATCACC	740 TAACATTAAAT ATGTGAATTA	760 CGAATATCTT GCTTATAGAA	780 TGGTCAAAT ACCAAGTTTA	800 TGGTTAACAC ACCAATTGTG	820 ATTGTACTTT TAACATGAAA	840 ATGTTGATCC TACAACCTAGG	860 AAAATGTATT TTTTACATA	880 GTAGATCTTT CATCTAGAAA
	TCTTTTGTAA AGAAAAACAT	820 TTCTCTTTAA AAGAGAAATT	840 GGAATAAGGT CCTTATTCCA	860 TTATCTAGTT AATAGATCAA	880 GATTTTGTAT CTAAAACTAC	900 GTTTATTGTA CAAATAACAT	920 GTGCTGGGAT CAGCACCCTA	940 AAGTTTCCAC TTCAAAGGTG	960 ATTGATACTC TAACATAGAG	980 GCCACACATT CGGTGTGTAA
	CTTCATTACT GAAGTAATGA	920 TAACTAATTG ATTGATTAAAC	940 GATATCGATT CTATAGCTAA	960 TTAACCCCTT AATTGGGAAA	980 TAATCGTAAT ATTAGCATTA	1000 TTGTTGTGTG AACAACACAC	1020 TTTATGACAC AAATACTGTG	1040 CATACAAGAT GTATGTTCTA	1060 ACATTATGTC TGTAATACAG	1080 TTACTGAGTG AATGACTCAC
	ACTCTTTGTT TGAGAAACAA	1020 GCTCTCTAAG CGAGAGATTG	1040 ATGTTGTAGT TACAACATCA	1060 TTGGATTCTC AACCTAAAGA	1080 TTGCTAAAGA AACGATTCTT	1100 AACTCAAAC TTGAGTTTGA	1120 ATAACTGATT TATTGACTAA	1140 TTACTGCTAC AATGACAGAT	1160 CATATATATG GTATATATAC	1180 TCAGTGGCCT AGTCACCGGA
	AGTAGGTTCA TCATCCAAGT	1120 TTAAGTAGAA AATTCTATCT	1140 ATCGGTGCGC TAGCCAGCGG	1160 AATTTTACTA TTAAAAATGAT	1180 ATTGGGAGAA TAACCCCTCT	1200 ACCACTAGAC TGGTGATCTG	1220 TACAACCAAA ATGTTGATCT	1240 TGTTCAATGA ACAAGTTACT	1260 CTTTAATAGT GAAATTATCA	1280 CTTCTGTTAT GAAGACAATA
	TTGTCGTGGA AACAGCACCT	1220 TATTTTAAAC ATAAAAAATTG	1240 CCCATGAAC GGGTACTTGA	1260 TTTGTATCTA AAACATAGAT	1280 GAAAAATCTC CTTTTATAGAG	1300 ATCCACTTCT TAGGTGAAGA	1320 CTTTTAGAAT GAAAACTTTA	1340 ACTTTGAATG TGAAACTTAC	1360 CGACTAAAAAG GCTGATTTTC	1380 TGAGTTTTTT ACTCAAAAAA
		1320		1340		1360		1380		1400

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100

TTTTCTAATA AAAAGATTAT	GACCTAAGAT CTGGATTCTA	AAAATCATCA TTTGTAGTAG	ATGGATAAGT TACCTATTCA	AGGAAATGGA TCCTTTACCT	AAGGTAACTC TTCCATTGAG	TTGTCACTAT AACAGTCATA	GTGTATATAT CACATATATA	ACAGCTCCTT TGTCGAGGAA	CTCATTTCTC GAGTAAAGGA
1420 TGATGTTGAC ACTACAACCTG	TCCATAAATG AGGTATTATTAC	1440 CTTGATCATG GAACTAGTAG	1460 AAAGCAAAAT TTTCGTTTAA	1480 TGTTAAATTT ACAATTTAAA	1500 GTAACCAACA CATTGGTGTG	1520 AAATGCACAG TTTACGTGTC	1540 ACTATAGACG TGATATCTGC	1560 AAGTATTAGG TTCATAATCC	1580 AACCGTATCT TTGGCATAGA
1600 ATCTGTCTCC TAGACAGAGG	1620 ATTTTACAAT TAAATGTGTA	1640 AGTCAAGCTC TCAGTTTCGAG	1660 TAGTTGTAGC ATCAACATCG	1680 TAGTTTCTTT ATTAAGTCTT	1700 ATTTAGTTCT TATACCTTAA	1720 CAAAGTGGCA ATATGGAATT	1740 CTATGCAAAAG GTTTCCACCG	1760 TGTTTTTAGT GATACGTTTC	1780 ACAAAAATCA
1800 TGAGATTAGT ACTCTAATCA	1820 CGTCTTATGC GCAGAATACG	1840 GTCTTACTAA CAGAATGATT	1860 TTGTTCAATT AACAAAGTAA	1880 TTTCTTCTTT AAAGAAGAAA	1900 TTGTGATTGA AACACTAACT	1920 TGTAATAATTA ACATTTTAAAT	1940 CTAAGTCACA GATTCAGTGT	1960 ACTTGAGATG TGAACCTCTAC	1980 TTACTAAAAA AATGATTTTT
2000 GATAAGAACG CTATTCTTGC	2020 TGTATAAATC ACATTATTGA	2040 GAAGTGAATT CTTCACCTTAA	2060 TGAAGCCAGT ACTTCGGTCA	2080 CTCTATTTCAT GAGATAAGTA	2100 ATCATAGCAT TAGTATCGTA	2120 TAATAGATCA ATTATCTAGT	2140 TGGACAACAC ACCTGTTGTG	2160 ATATATAGGA TATATATCCT	2180 TTAGAGCTGT AATCTCGACA
2200 CATGACCTTC GTACTGGAAG	2220 CCGGAAATGC GGCCTTTTACG	2240 TAAATCAGTT ATTTAGTCAA	2260 TCCTGGTTTA AGAACCAAAAT	2280 TCCTTTTGGG AGGAAAAAACC	2300 AGTATCATGA TCATAGTACT	2320 TATCATTTAG ATAGTAAATC	2340 CCAAAGGTTT GGTTTCCAAA	2360 TTGGTTTCAG AACCAGGATC	2380 TATTCGGATT ATAAGGCTAA
2400 CGTTTGACGT GCAAACTGCA	2420 TATGTGTGAA ATACACACTT	2440 AGCGTCAATA TCGCAGTTAT	2460 ACTAAACCTT TGATTTTGAA	2480 GGATTGACTA CCTAACTGAT	2500 GTCAAAATAT CAGTTTATA	2520 AAACTGATTG TTTGACTAAC	2540 CATTGAATTC GTAACCTAAG	2560 TTGAAAAATT AACTTTTAAA	2580 TCCCTTAAAA AGGGAATTTT
2600 TGAACATGAA ACTTGTACTT	2620 TTTCATCAAG AAAGTAGTTC	2640 ATTTTGTCTT TAAACACAGAA	2660 TTGGAAGGAT AACCTTCCTA	2680 GTGATTTATA CACTAAATAT	2700 ATCTATACAA TAGATATGTT	2720 TCATACATTT AGTATGTAAA	2740 TGCATGATAT ACGTACTATA	2760 TAGTTTTTTG ATCAAAAAAC	2780 AAGAACCAAA TTCTTGGTTT
2800 AATAGAGCTT TTATCTCGAA	2820 CTTTATAAAA GAAATATTTT	2840 CTGATTTAGC GACTAAATCG	2860 CTTGATAAGA GAACCTATCT	2880 AAAAGAAGGT TTTCTTCCCA	2900 AGATAATCGA TCTATTAGCT	2920 ACTCATGGGG TGAGTACCCC	2940 ATGAGTTAAA TACTCAATTT	2960 AATGTGTGCA TTACACACGT	2980 CTTAGTTTCT GAATCAAGA
3000 AAAACCTTTT TTTTGGAAAA	3020 GAAGTCGAAA CTTCAGCTTT	3040 CAATGACAAT GTTACTGTTA	3060 ATTGGCTGCG TAACCGACGC	3080 AAGTTGATAT TTCAACTATA	3100 ATAACAGGAT TATTGTCTTA	3120 CTTAAAGTTG GAATTTCAAC	3140 AAATTTGATA TTTAACTATT	3160 TTCAGATTTT AAGTCTAAAA	3180 AATTTTAGAG TTAAATCTCT
3200 CACCAGATGA GTGGTCTACT	3220 TCAGAGTTTC AGTCTCAAAG	3240 AGATTTACAT TCTAAATGTA	3260 TTGAAGTATA AATCTCATAT	3280 AAACATTTTG TTTGTAAAC	3300 AACACATATA TTGTGTATAT	3320 TCTAAAGCAG AGATTTCTGC	3340 TAACCTCAAA ATTGAAGTTT	3360 AATAGGGTAA TTATCCCATT	3380 CTAATAGTAA GATTATCATT
3400 CTTACATTTG GAATGTAACA	3420 TTTTTTTAAAT AAAAAAATTA	3440 GCTTTTATAC CGAAAAATATG	3460 TTACTATCAT AATGATAGTA	3480 TTTTATATAT AAAAATATATA	3500 AGATGCTGTG TCTACGGACC	3520 TTAAGTAAAG AATTCATTTT	3540 ATGATTATCA TACTAATAGT	3560 AAAACCTGTT TTTTTGACAAC	3580 TTTGTAGTAA CAATCATTTG
3600 GAAATTTGTT CTTTAACAC	3620 CAAAATGTAAC GTTTACATTG	3640 ATATTATATA TATAATATAT	3660 AGCTTTTCTT TCGAAAGAAA	3680 CACTTTGGTG GTGAAACCAC	3700 CATTCTCTCT GTAAGAGAGA	3720 AAATAATGGC TTTATTACCG	3740 CTCTATTGAT GAGATACTA	3760 CGAGTATCTG CGTCATAGAC	3780 ATTCCTAGTT TAAGATCAAA
3800 TTGAAATGGT AACCTTTACCA	3820 TTTTGTCATA AAAACGTAAT	3840 ATTATTGTTT TAATGCAATT	3860 TAATGCAATT ATTACGTAAA	3880 TTGTTTTATC AACAATAAG	3900 TCCAGCATTT AGGTCGTAAA	3920 CCAAGATCAT GGTTCGTAGT	3940 TGATCGTTAT ACTAGCAATA	3960 GAAATACAAC CTTTATGTTG	3980 ATGCTGATGA TACGACTACT
4000 ACTTAGAGCC TGAATCTCGG	4020 TTAGTAAGTA AATCATTCAT	4040 ATTAGCTAAG TAATCGATTG	4060 AACGTCATTC TTGCAGTAAG	4080 TAATATTCTT ATTATAAGAA	4100 CTGGATCGCG GACCTACGCC	4120 TTTTTGGTGT AATAAACCACA	4140 TATGAAGGAT ATACTTCTCT	4160 AGAAGCGCTG TCTTCGCGAC	4180 TTCAAGCCGG AAGTTCGGCC
4200 AGAAACCTCA TCTTTGGAGT	4220 ATGTTTTGAA TACAAAACCT	4240 CTCGTAACAC GAGCATTTGT	4260 CGAACTTAAT GCTTGAATTA	4280 TCTCTAGAGT AGAGATCTCA	4300 TACAGTTATT ATGTCAATAA	4320 GTGTCTACTG CACAGATGAC	4340 GAAAAATACAA CTTTTATGTT	4360 GAACCTCACA CTTGAAGTGT	4380 ATCTTTCTGA TAGAAAGACT
4400 CCATTCCCTT GGTAAGGAAA	4420 TCTTCATGTG AGAAGTACAC	4440 CAGGATCTTG GTCCTAGAAC	4460 AAGAAAAAAT TTCTTTTTTA	4480 TCAGAAATTA AGTCTTAATA	4500 CTTCCACACA GAAGGTGTGT	4520 AGGAGTTACT TCCTCAATGA	4540 AGAAACAGTC TCTTTGTCTG	4560 CAAAGTTAG GTTTCCAATC	4580 CAGTACGACA GTCATGCTGT
4600 CATTTTTCTC GTAAAAAGAG	4620 CCCTCTTCTT GGGAGAAGAA	4640 CTGATAAAAA GACTATTTTT	4660 AAATGTTTTT TTTACAAAAA	4680 TTTCTTTTGT AAAGAAAACA	4700 CTACTTTGTA GATGAACACT	4720 ATACAGCAAG TATGTCGTTC	4740 CTTGAAGAAC GAACCTCTTG	4760 CAAATGTCGA GTTTACAGCT	4780 TAATGTAAGT ATTACATTTA
4800 GTAGATTCTC CATCTAAGAG	4820 TAATTTCTCT ATTAAGAGAGA	4840 GGAGGAACAA CCTCCTTGTG	4860 CTTGAGACTG GAACTCTGAC	4880 CTCTGTCCGT GAGACAGGCA	4900 AAGTAGAGCT TTCATCTCGA	4920 AGGAAGGTAT TCCTTCCATA	4940 ATGTGCTGCT TACACGACGA	4960 ACTAAGTATG TGATTTCACTA	4980 TCAACCAATT AGTTGGTTAA
5000 ACTCCACAAA TGAGGTGTTT	5020 ACCTTCTTTT TGGAAGAAAA	5040 TAGTTAGTTA ATCAATCAAT	5060 TCCTAGAACA AGGATCTTGT	5080 ATCTTTTGAC TAGAAAACTG	5100 ATAAATCTTA TATTTAGAAT	5120 ATGTCTTGTG TACAGAACAA	5140 ATAGGCAGAA TATCCGTCTT	5160 CTGATGATGG GACTACTACC	5180 AGTATATCGA TCATATAGCT
5200 GTCCCTTAAA CAGGGAAATT	5220 GAAAAGGTTA CTTTTCCAAT	5240 GTGCTTTGGT CACGAAACCA	5260 TTTTATTTC AAAATAAAAG	5280 GATAAAGGCC CTATTTCGGG	5300 ATATTCTAGG TATAAGATCC	5320 CTATGATGAT GATACTACTA	5340 TCTTGAATTC AGAACTTAAAG	5360 TATTAAACCTG ATAATTGGAC	5380 CTGAGTCTAC GACTCAGATG
5400 AGATTACTAT TCTAATGATA	5420 ATATATATAT TATATAGAAA	5440 ATATATCTTT TATATAGAAA	5460 TGGTCTTGTC ACCAGAACAG	5480 TTAGTTCCCTG AATCAAGGAC	5500 ATTTAGTATT TAAATCATAA	5520 GGCTTCATTC CCGAAGTAAG	5540 AGGTGAAACC TCCACTTTGG	5560 CTAATGAGAA GATTACTCTT	5580 TTAAAAAAC AATTTTTTTG
5600 AAGCAGTTTT TTCGTCAAAA	5620 AAACTCTTGA TTTGAGAACT	5640 TCAAAATCCA AGTTTAGGTT	5660 CCTTCCCTC GGAAAGGGAG	5680 ATAAAGTGTC TATTTACAG	5700 GAATTTGGAT CTTAAACCTA	5720 GAGGATGATT CTCCTACTAA	5740 TATGTTTCGA ATACAAGGCT	5760 GAAGGAAAAA CTTCCTTTGT	5780 TGTTTGGAAA ACAAACCTTT
5800 TAGCTATAGA ATCGATATCT	5820 AGTTGTTAGA TCAACAATCT	5840 AACTAATGAC TTGATTACTG	5860 CTTATGATCT GAATACTAGA	5880 TTTCCAAACA AAAGGTTTGT	5900 GGAGAAATTG CCTCTTTAAC	5920 CTGAGAGAAG GACTCTCTTC	5940 AGAACCAGGT TCTTGGTCCA	5960 TCTGGCTAGC AGACCGATCG	5980 CAGGTAACAA GTCCATTGTT
6000 TGACCACAAT ACTGGTGTTA	6020 ATCTTCTGCT TAGAAGACGA	6040 CTTGAAGCTA GAACTTCGAT	6060 ATTAACTACT TAATTAGTGA	6080 TTATACGTCC AATATGACAG	6100 CCGTTATAGA GGCAATATCT	6120 GAGATACACA CTCTATGTGT	6140 TATACACGTA ATATGTGCAT	6160 CATGAAAACT GTACTTTTGA	6180 AAAAGTTGAA TTTTCAACTT
6200 GGACTTTGAT CCTGAAACTA	6220 GGATACTAGA CCTATGATCT	6240 CAATTATAGT GTAAATATCA	6260 GAAACCTTAA CTTTGGGATT	6280 ATATGTGATA TATACACTAT	6300 AGTGATAACA TCACATATGT	6320 AAATGCTTTT TTTACGAAAA	6340 AAAATCTATC TTTTAGATAG	6360 TTTCTTGTTA AAAGAACAAT	6380 ATTTAGTAGC TAAATCATCG

3920 3940 3960 3980 4000
 TGTGAGAGAA GAAAGGTATG TCTCACCGAT GAAAGATACT CAAAACCCGG TATTTTAAAT TTGTGAAATT TGCAAATAAA AAAAATGCTT TCTACAAGAT
 ACAGTCTCTT CTTTCCATAC AGAGTGGCTA CTTTCTATGA GTTTTGGGCC ATAAAAATTA AACACTTTAA ACGTTTATTT TTTTACGAA AGATGTTCTA
 4020 4040 4060 4080 4100
 AGATTAATTT CTTGCAATGT TTAGTAGCTG TAGAAAAAAA AGAAATGTAA GAAAGTTTCT TACAGATGGG AAAGAATACG TTGCTGGCAA CAGATGATGA
 TCTAATTAAG GAACGTTACA AATCATCGAC ATCTTTTTTT TCTTTACATT CTTTCAAAGA ATGTCTACCC TTTCTTATGC AACGACCGTT GTCTACTACT
 4120 4140 4160 4180 4200
 GAGAGGAATG TTTCCGGGAA GTAGCTCCGG CAACAAAATA CCGGAGACTC TCCCGCTGCT CAATTAGCCA CCATCATCAA CGGCTGAGTT TTCACCTTAA
 CTCTCCTTAC AAAGGCCCTT CATCGAGGCC GTTGTTTTAT GGCCTCTGAG AGGGCGACGA GTTAATCGGT GGTAGTAGTT GCCGACTCAA AAGTGAATT

SEQ ID NO:48 and SEQ ID NO:49

Alternatively splice Arabidopsis AGL27 cDNA and resulting Alternate Arabidopsis AGL27 amino acid sequence

20 40 60 80 100
 ATGGGAAGAA GAAAAATCGA GATCAAGCGA ATCGAGAACA AAAGCAGTCG ACAAGTCACT TTCTCCAAAC GACGCAATGG TCTCATCGAC AAAGCTCGAC
 TACCCCTCTT CTTTTCAGCT CTAGTTCGCT TAGCTCTTGT TTTTCGTCAGC TGTTCACTGA AAGAGGTTTG CTGCGTTACC AGAGTAGCTG TTTTCGAGCTG
 M G R R K I E I K R I E N K S S R Q V T F S K R R N G L I D K A R>
 120 140 160 180 200
 AACTTTCGAT TCTCTGTGAA TCCTCCGTCG CTGTTGTCGT CGTATCTGCC TCCGGAAGAA TCTATGACTC TTCCTCCGGT GACGAGATAG AAGCGCTGTT
 TTGAAAGCTA AGAGACACTT AGGAGGCGAG GACAACAGCA GCATAGACGG AGGCCTTTTG AGATACTGAG AAGGAGGCCA CTGCTCTATC TTGCGGACAA
 Q L S I L C E S S V A V V V V S A S G K L Y D S S S G D E I E A L F>
 220 240 260 280 300
 CAAGCCGGAG AAACCTCAAT GTTTTGAACG CGATCTTGAA GAAAAAATTC AGAATTATCT TCCACACAAG GAGTTACTAG AAACAGTCCA AAGCAAGCTT
 GTTCGGCCTC TTTGGAGTTA CAAAACCTGA GCTAGAACTT CTTTCTTAAG TCTTAATAGA AGGTGTGTTT CTCAATGATC TTTGTGAGGT TTCGTTTCGAA
 K P E K P Q C F E L D L E E K I Q N Y L P H K E L L E T V Q S K L>
 320 340 360 380 400
 GAAGAACCAA ATGTGATGAA TGTAAGTGTA GATTCTCTAA TTTCTCTGGA GGAACAACCT GAGACTGCTC TGTCGTAAG TAGAGCTAGG AAGGCAGAAC
 CTTCTTGGTT TACAGCTATT ACATTCACAT CTAAGAGATT AAAGAGACCT CTTTGTGTAAG CTCTGACGAG ACAGGCATTC ATCTCGATCC TTCGCTCTTG
 E E P N V D N V S V D S L I S L E E Q L E T A L S V S R A R K A E>
 420 440 460 480 500
 TGATGATGGA GTATATCGAG TCCCTTAAAG AAAAGGAGAA ATTGCTGAGA GAAGAGAACC AGTTTCTGGC TAGCCAGATG GGAAAGAATA CGTTGCTGGC
 ACTACTACCT CATATAGCTC AGGGAATTTT TTTTCTCTT TAACGACTCT CTTCTCTTGG TCCAAGACCG ATCGGCTCTAC CCTTTCTTAT GCAACGACCG
 L M M E Y I E S L K E K E K L L R E E N Q V L A S Q M G K N T L L A>
 520 540 560
 AACAGATGAT GAGAGAGGAA TGTTCCTGGG AAGTAGCTCC GGCAACAAAA TACCGGAGAC TCTCCCGCTG CTCAATTAG
 TTGTCTACTA CTCTCTCCTT ACAAAGGCC TTCTATCGAG CCGTTGTTTT ATGGCCTCTG AGAGGGCGAC GAGTTAATC
 T D D E R G M F P G S S S G N K I P E T L P L L N *>